HUAWEI CLOUD MaaS Security White Paper

Issue 1.0

Date 2026-09-10

HUAWEI CLOUD COMPUTING TECHNOLOGIES CO.,LTD.

Steering Committee

Hu Yuhai President, Public Cloud Service Department, HUAWEI CLOUD

Zeng Yubin Chief Cybersecurity and Privacy Protection Officer, HUAWEI CLOUD

Lead Authoring Group

ZOU FENG, Jia Yingjie

Key Members of the Authoring Group

Yao Huihui, Ling Dongdong, He Yingqin, Jin Peisong, Huang Jianqiang, Yao Yingquan, Chen Yanhui, Ni Jianqiang, Chen Li, Chen Long

Special Acknowledgements

Huang Yanxia, Yang Yan, Shu Yanfeng, Guo Jia, Wu Jun, Zhai Minggang, Chen Changqing, Chen Fangfei, Chen Jiwei, Dang Zhongxing, Rao Yixuan, Huang Yitian, Yin Aijun, Zhang Zhongyong

HUAWEI CLOUD Computing Technologies Co., Ltd.

Address: HUAWEI CLOUD Data Center Jiaoxinggong Road
Qianzhong Avenue
Gui'an New District
Gui Zhou 550029
People's Republic of China

Website: **https://www.huaweicloud.com/intl/en-us/**

Content

Executive Summary

As large-model capabilities continue to mature, enterprises are shifting their AI adoption from self-built models and ad-hoc pilots toward platform-centric, service-oriented and large-scale deployment. MaaS (Model as a Service) encapsulates model provisioning and invocation, model development, deployment and inference, as well as security governance and operational capabilities into platform-based services, enabling enterprises to rapidly select, deploy and consume models in response to business demands. Built on cloud infrastructure and a model service platform, HUAWEI CLOUD MaaS delivers an integrated hosting solution spanning computing power, models, development tools and lifecycle management.

As MaaS moves into production environments, the scope of security protection expands beyond standalone models to cover the full chain including computing infrastructure, model services, data, Agents, and operations & maintenance. Centered on core security challenges arising from MaaS adoption, this whitepaper systematically describes HUAWEI CLOUD' s MaaS security governance framework, security responsibility demarcation, platform-side security safeguards, together with security controls and configuration best practices for clients to securely consume MaaS.

Scope of Application

This whitepaper is intended for enterprises and organizations planning to rapidly embed large-model capabilities into existing business workflows. It serves as a reference for AI and MaaS architects, model and Agent application developers, cybersecurity, data security and AI security practitioners, platform operation and maintenance staff, as well as professionals in charge of data governance, privacy protection, risk management and compliance review within the organization.

Key Content of the Whitepaper

Based on the shared-responsibility model — **where HUAWEI CLOUD secures the cloud and managed MaaS services themselves, while clients take responsibility for the security of content they configure, upload, develop and consume on the MaaS platform** — this whitepaper addresses security requirements throughout the build, usage and continuous operation phases of MaaS, and focuses on answering the following questions:

First, what security risks MaaS faces. It analyzes changes to conventional risks and AI-specific risks introduced by scenarios such as model-as-a-service delivery, multi-tenant computing, RAG and Agents.

Second, how to build a systematic MaaS security regime. Guided by AI governance principles, management systems and end-to-end lifecycle controls, a security assurance framework is established covering computing infrastructure, models, data, Agents, and operations and maintenance.

Third, the respective responsibilities of HUAWEI CLOUD and clients. The security responsibility boundary is defined according to the actual scope of control exercised by each party over infrastructure, platforms, models, data and applications.

Fourth, how HUAWEI CLOUD safeguards the MaaS platform. It describes the security controls deployed by HUAWEI CLOUD for the computing foundation, model services, in-platform data transmission and storage, ongoing security operations, and baseline security of the Agent platform.

Fifth, how clients can use MaaS securely. It specifies mandatory security requirements for clients covering model selection and consumption, data lifecycle management, operations and maintenance, and Agent applications, together with actionable best-practice security configurations for reference.

Basic Security Requirements for Clients Using MaaS

Before adopting MaaS for production-grade business operations, clients are recommended to complete at least the following security configurations:

Secure connectivity: Select appropriate access methods for model services based on the business architecture to reduce unnecessary exposure to the public network.

Identity and access control: Manage accounts, roles, resource permissions and API Keys in accordance with the principle of least privilege.

Data protection: Clarify data sources, authorizations, sensitivity levels and intended purposes, and enforce data minimization, data classification and grading, as well as sensitive information protection.

Model and application security: Select suitable models for business scenarios, and implement necessary security controls for input and output, RAG retrieval, and Agent tool invocation.

Operations and audit: Establish a continuous monitoring and response mechanism for AI assets, logs, abnormal invocations and security incidents.

With the above controls in place, clients can build a closed-loop security workflow for MaaS usage covering “secure connectivity — identity authorization — data protection — model defense — asset management — monitoring and audit — incident response”. HUAWEI CLOUD commits that with respect to client content, **except for processing necessary to provide the platform and related services to the client, HUAWEI CLOUD will not use client content for purposes including model training, fine-tuning or optimization without the client’s explicit authorization.**

1 AI Business Trends and Risk Challenges

1.1 AI Business Development Trends

As the capabilities of large models continue to mature, enterprises are gradually shifting their AI adoption from point-based pilots and self-built models toward platform-centric, service-oriented and large-scale applications. The traditional self-building model generally requires enterprises to independently undertake computing infrastructure construction, model selection and training, model deployment, security protection and continuous operations. It features high investment costs, long implementation cycles, and imposes stringent requirements on professional talents as well as security and compliance capabilities.

MaaS (Model as a Service) encapsulates model provisioning and invocation, model development, model deployment and inference, together with security governance and operational capabilities into platform-based services, enabling enterprises to select, deploy and invoke models according to business requirements. Built on cloud infrastructure and the model service platform, HUAWEI CLOUD MaaS provides clients with an integrated hosting solution covering computing power, models, development tools and lifecycle management. It helps clients lower the barriers to building and adopting AI capabilities and accelerates the integration of AI applications into real-world business workflows.

Against the large-scale adoption of MaaS, enterprise AI businesses are undergoing four major shifts:

- **First, the model consumption paradigm is evolving from "self-built for self-use" to "on-demand acquisition":**

Enterprises can access model capabilities via model invocation, a model marketplace, dedicated deployment and online inference. Models are gradually turning into composable, configurable and sustainably operated cloud services.

- **Second, deployment architectures are shifting from standalone environments toward the collaboration of cloud, edge, endpoint and hybrid frameworks:**

Computing tasks, model weights and inference requests may flow across different regions, computing resource pools and business environments. Enterprises need to concurrently address service performance, data flows, resource isolation and cross-environment security.

- **Third, application patterns are expanding from content generation to task execution:**

Enterprises are connecting models with knowledge bases, business systems, APIs, plugins and MCP services to further build Agents. Instead of merely answering questions, AI applications are moving toward task planning, tool invocation and business workflow execution. Controllability over permission scopes, behavioral boundaries and execution outcomes has become far more critical.

- **Fourth, AI applications are transitioning from partial pilots to integration with core business operations:**

MaaS is enabling AI capabilities for scenarios such as office work, R&D, client service, finance, manufacturing, healthcare and government services. Once model services are embedded into critical business workflows, security incidents will affect not only model outputs, but also business continuity, data compliance, client rights and corporate reputation.

Accordingly, the scope of security protection for MaaS is no longer limited to individual models. It extends to the full service chain covering computing infrastructure, model services, Agent applications, client data as well as operation and maintenance activities. As the MaaS service provider, HUAWEI CLOUD shall secure cloud infrastructure, MaaS models and the platform within its scope of responsibility. Clients, in turn, bear corresponding responsibilities for the data they upload and consume, account privileges, business configurations, application security and operational behaviors.

1.2 1.1 Security Risks and Compliance Challenges in the AI Context

1.2.1 MaaS Security Risks

While the widespread adoption of AI drives efficiency gains, it also introduces risks that differ markedly from those of traditional IT security. The OWASP Top 10 for Large Language Model Applications (2025) outlines core risk categories for LLM-powered applications, including prompt injection, adversarial model attacks, data leakage, supply-chain vulnerabilities, excessive plugin privileges and unauthorized Agent execution. Unlike traditional web security, which mainly focuses on access control, injection flaws, misconfigurations and component vulnerabilities, AI security differs primarily in the following aspects:

- Attack targets expand from systems and code to models and semantic interaction processes. Attackers do not necessarily need to breach conventional system perimeters. Instead, they may manipulate model behavior via prompts, context, knowledge bases or external documents, compelling the model to generate unintended content, disclose sensitive information or trigger improper tool invocations.
- Attack methodologies evolve beyond syntactic exploitation to incorporate deeply layered semantic manipulation. Prompt injection leverages the model’s tendency to follow natural-language instructions to bypass guardrails or execute unapproved tasks. Conventional rule-based pattern-matching techniques alone cannot deliver reliable blocking.
- The nature of vulnerabilities shifts from deterministic code defects to highly context-dependent behavioral deviations of models. Influenced by the dynamic

interplay of system prompts, retrieved augmented content and conversational context, models may produce unexpected outputs for similar inputs in the absence of strict constraints or amid contextual drift — outcomes that traditional rules struggle to cover.

- The scope of protected assets extends beyond source code and databases to encompass model weights, system prompts, training and fine-tuning datasets, enterprise vector knowledge bases, inference logs, plugin configurations and Agent execution trails.
- Risk impacts go beyond technical consequences such as data leakage and service disruption, extending to flawed business decisions, privilege escalation and downstream system risks.

Overall, AI security risks are characterized by semantic-driven attacks, probabilistic behavioral outcomes, model-centric assets and business-oriented impacts. In addition to conventional security controls, MaaS platforms must build dedicated capabilities targeting model behavior, RAG retrieval, Agent execution and security operations.

1.2.2 AI Compliance Challenges

As artificial intelligence technologies are deeply integrated into enterprise business workflows and industrial production systems, the resulting compliance challenges have expanded beyond traditional cybersecurity, data security and personal information protection to cover model governance, content safety, algorithm transparency, risk classification, human oversight, audit traceability and liability allocation.

Worldwide, AI regulation is evolving from principled initiatives and ethical guidelines toward binding legal regimes and enforceable compliance obligations. While jurisdictions differ in regulatory stringency, legislative models and implementation timelines, they share a common overall direction. By adopting mechanisms such as-based risk classification, transparency mandates, entity accountability, human oversight and continuous assessment, governments embed AI risk governance throughout the full lifecycle of product design, development, deployment, operation and decommissioning.

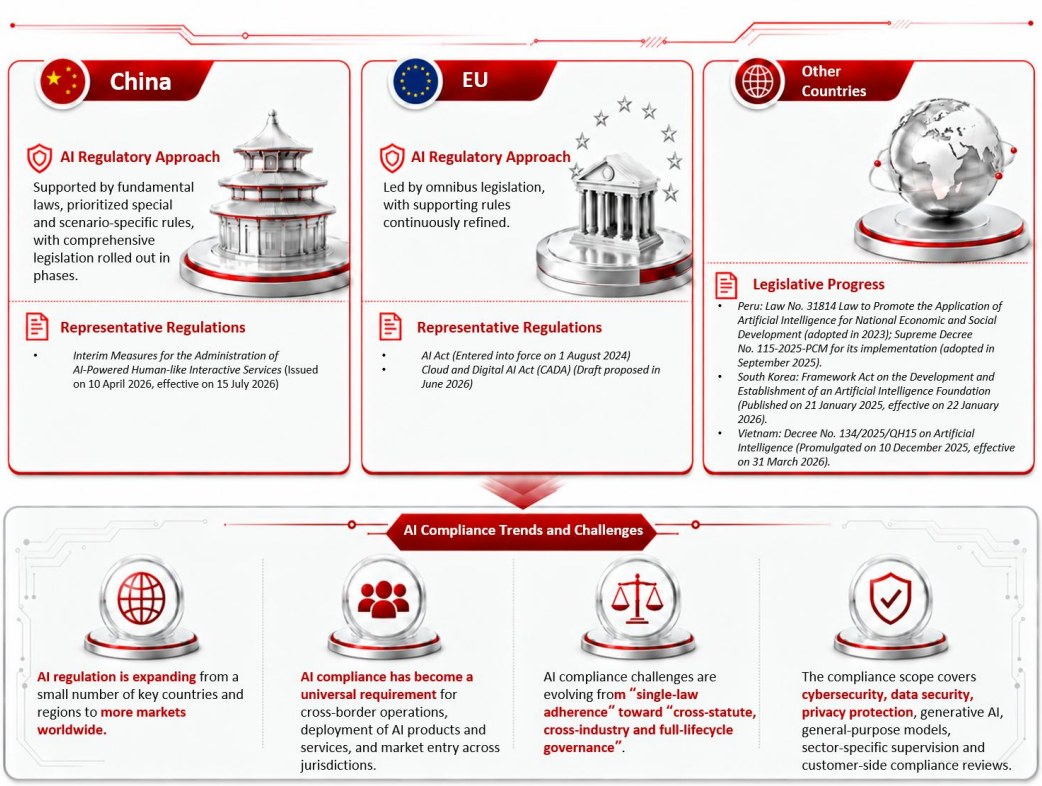

Figure 1-1 AI Compliance Trends and Challenges

China currently adopts a governance path of **fundamental legal support, priority for special and scenario-specific rules, and phased advancement of comprehensive legislation**. At this stage, China's AI regulatory system is built upon the *Cybersecurity Law*, *Data Security Law* and *Personal Information Protection Law*. Specialized regulatory rules have gradually taken shape for algorithmic recommendation, deep synthesis, generative AI services and other domains.

The *Interim Measures for the Administration of Generative Artificial Intelligence Services* clarify the obligations of generative AI service providers covering training data, generated content security, user right protection, complaint handling, security assessment and algorithm filing. Where services with public opinion-shaping or social mobilization capabilities are offered to the public within China, providers shall conduct statutory security assessments and complete formalities for the filing, modification and cancellation of algorithms and large-scale models. The *Interim Measures for the Administration of AI-Powered Human-like Interactive Services*, issued on 10 April 2026 and effective on 15 July 2026, further extend regulation to human-like AI services featuring continuous emotional interaction. Providers are required to establish a full-lifecycle security management system covering deployment, operation, upgrade and decommissioning, with a focus on mitigating risks such as over-reliance and emotional manipulation. Enhanced governance is mandated for minors, the elderly, sensitive personal data, interaction logs, security assessments and algorithm filings.

For industrial development, the *Opinions on Further Implementing the "AI+" Initiative* released by the State Council promote deep integration of artificial intelligence across economic and social sectors. It highlights security and controllability, differentiated sector-specific policies, and open sharing. While this document does not serve as a direct basis for administrative penalties, it provides critical guidance for industry-wide compliance construction, regulatory inspections and scenario-based governance. It also signals that future AI governance will place greater emphasis on application scenarios, sectoral characteristics, data security and technical controllability.

Enterprises operating in heavily regulated sectors must additionally comply with industry-specific requirements. For instance, the *Guiding Opinions on the Secure Development and Application of Artificial Intelligence in the Banking and Insurance Sectors*, published by the State Administration of Financial Regulation, sets requirements for governance frameworks, development and deployment, data governance, computing infrastructure and risk management. It mandates banking and insurance institutions to build a full-lifecycle AI management system and incorporate AI-related risks into their enterprise-wide risk management frameworks.

Beyond general-purpose AI regulation, stricter superimposed controls apply to high-supervision industries. Taking the financial sector as an example, the above-mentioned guiding opinions issued by the State Administration of Financial Regulation define requirements for governance architecture, development and application, data governance, computing power construction, risk management, capability improvement, assurance and oversight. Banking and insurance institutions are required to implement an end-to-end governance regime for AI and embed AI application risks into corporate risk management systems.

The European Union follows a regulatory strategy of **flagship omnibus legislation first, with supporting rules refined on an ongoing basis**. The *EU AI Act* entered into force on 1 August 2024, with obligations to apply in staggered phases. It adopts a risk-based classification regime that imposes differentiated controls on prohibited-use AI, high-risk AI, AI subject to specific transparency obligations, and low- or minimal-risk AI. It also lays down requirements for general-purpose AI models regarding transparency, copyright protection, model evaluation, security safeguards and systemic risk governance.

Following the establishment of the foundational regulatory framework under the AI Act, the EU is refining detailed rules for general-purpose AI, generative content labelling, high-risk AI systems and model security assessments through implementation guidelines, codes of practice, technical standards and regulatory instruments. These measures transform the principled provisions of the AI Act into enforceable, auditable concrete rules. Meanwhile, the European Commission tabled the draft *Cloud and AI Development Act (CADA)* in June 2026. The proposal aims to boost the underlying capacity, competitiveness and technological sovereignty of the EU's AI industry by accelerating the build-out of cloud, computing and data-center infrastructure and introducing a cloud-AI sovereignty evaluation framework.

AI legislation is rapidly expanding from key jurisdictions including China and the EU to an increasing number of countries. Peru enacted Law No. 31814, the *Law to Promote the Application of Artificial Intelligence for National Economic and Social Development*, in 2023, with its implementing regulation adopted via Supreme Decree No. 115-2025-PCM in September 2025. South Korea promulgated the *Framework Act on the Development and Establishment of an Artificial Intelligence Foundation* on 21 January 2025, which took effect on 22 January 2026. It establishes foundational governance for AI industrial development, transparency, high-impact AI and the security management of generative AI. Vietnam issued the *Artificial Intelligence Law* (Decree No. 134/2025/QH15) on 10 December 2025, effective 1 March 2026, marking a shift from policy-driven AI strategies to dedicated statutory regulation.

These legislative developments demonstrate that **AI oversight is extending from a small set of priority jurisdictions to global markets. AI compliance is no longer a niche concern exclusive to large technology firms; it has become a universal requirement for enterprises conducting cross-border operations, deploying AI products and services, and entering overseas markets.**

Overall, AI compliance challenges are evolving from **single-statute compliance toward cross-jurisdictional, cross-industry and full-lifecycle governance**. AI compliance programmers must satisfy baseline cybersecurity, data security and privacy obligations while

also supporting regulatory controls for generative AI services, governance of general-purpose models, vertical-sector supervision and client-side compliance reviews. Accordingly, MaaS security capabilities should not be limited to model invocation protection and platform defenses. They need to be further built into configurable risk controls, traceable log auditing, verifiable security evidence and deliverable compliance artefacts for clients.

2 MaaS Security Governance Framework

MaaS connects cloud infrastructure, model services, data processing and business applications into an end-to-end service chain. Its security governance not only protects models themselves, but also coordinates multiple components including computing infrastructure, model provisioning and runtime, client data, Agents and applications, as well as ongoing operations and maintenance, while clarifying the security responsibility boundaries among HUAWEI CLOUD, clients and relevant third parties.

HUAWEI CLOUD MaaS security governance is built upon the AI security governance framework, delivering unified methodologies and requirements across governance principles, management systems and the full lifecycle of AI systems. Tailored to the platform-centric, service-oriented and multi-stakeholder collaborative nature of MaaS, MaaS-specific security governance translates these requirements into practical activities covering MaaS platform planning, design and development, testing and validation, release and deployment, operational monitoring, and decommissioning. It establishes systematic security capabilities spanning computing infrastructure, models, Agents, data, and operations and maintenance.

Therefore, the AI security governance framework illustrated in the figure below serves as the foundational framework and methodological basis for HUAWEI CLOUD to implement MaaS security governance. By integrating this framework with real-world MaaS service scenarios, HUAWEI CLOUD has established a MaaS security governance system covering governance and management, the full lifecycle of AI systems, and AI security capabilities, providing unified guidance for the secure, trustworthy and sustainable operation of MaaS services.

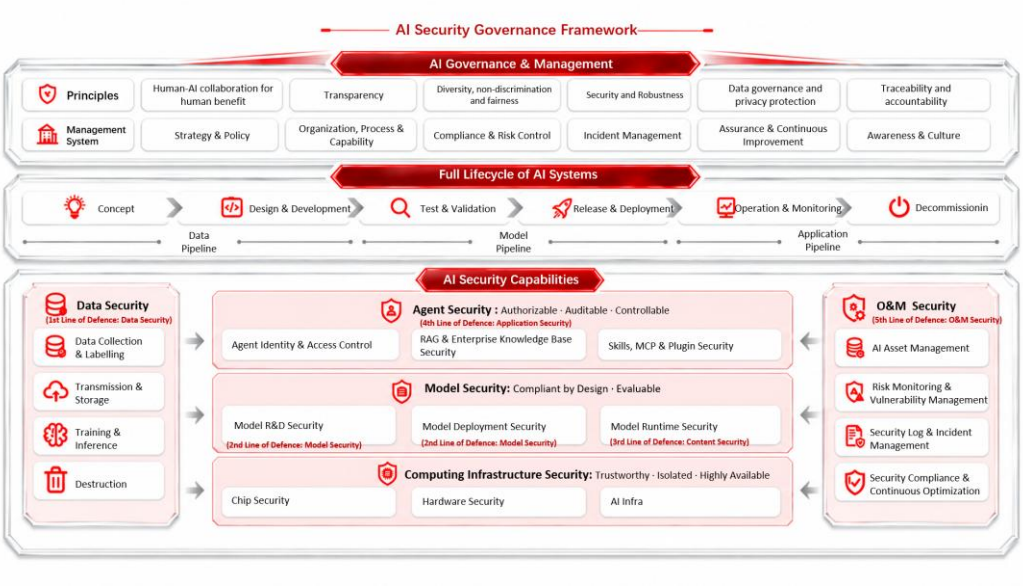

Figure 2-1 AI security governance framework

The Governance & Management Layer defines security objectives and management mechanisms

Six AI governance principles specify the baseline requirements for HUAWEI CLOUD AI development and adoption. Six management elements — Strategy & Policy, Organization-Process-Capability, Compliance & Risk Control, Incident Management, Assurance & Continuous Improvement, and Awareness & Culture — translate high-level principles into enforceable organizational responsibilities, policies and management activities.

The Full-Lifecycle AI System Layer defines the implementation workflow for security requirements

Security controls run through concept, design & development, test & validation, release & deployment, operation & monitoring and decommissioning phases. Embedded into real-world R&D and operation workflows via the data pipeline, model pipeline and application pipeline, these controls enable early risk identification, in-process governance and continuous improvement.

The AI Security Capability Layer delivers technical and operational safeguards tailored for MaaS

Three vertical capabilities are built for computing infrastructure, models and Agents respectively: trustworthiness-isolation-high-availability for the underlying compute layer, controllability-compliance-evaluability for models, and authorization-audit-governance for Agents. Data security and O&M security serve as two horizontal cross-cutting capabilities across all layers.

To address the feature that AI risks propagate and escalate across data, model, application and operational stages, HUAWEI CLOUD establishes five interlocking lines of defence: data security, model security, content security, application security, and O&M security. Working in tandem, these five lines form a defence-in-depth system covering source-side prevention, in-flight control, runtime monitoring and closed-loop remediation.

2.1 AI Governance Principles

HUAWEI CLOUD adheres to the six AI governance principles and translates them into enforceable security requirements throughout the design, development, operation of MaaS services and client usage lifecycle.

Figure 2-2 AI Governance Principles

- **Human-AI Collaboration for Human Benefit:** AI shall operate under human authorization and oversight, collaborate with humans by complementing their strengths, serve defined objectives without undermining human autonomy.
- **Transparency:** AI applications shall safeguard users' right to know and right to choose, maintain comprehensibility and explainability, enabling users to understand the rationale, accuracy and limitations behind AI decisions.
- **Diversity, Non-discrimination and Fairness:** The design, training and deployment of AI shall prevent bias and discrimination, ensure diverse and high-quality datasets, and improve model explainability.
- **Security and Robustness:** AI systems shall mitigate risks to an acceptable level in case of failures or anomalies, with secure, reliable and resilient protection capabilities.
- **Data Governance and Privacy Protection:** AI systems shall be developed based on high-quality, representative data, while protecting user privacy and associated rights across the full lifecycle.
- **Traceability and Accountability:** AI systems shall define clear responsible entities, incorporate evaluable and traceable mechanisms, and manage risks and liabilities through tiered, role-based governance.

2.2 AI Management System

To drive the sustained implementation of AI governance principles within MaaS services, HUAWEI CLOUD has established a management system covering decision-making, execution, oversight and improvement. Centered on the six management elements, this system translates governance principles into organizational responsibilities, management policies, R&D workflows, risk controls and ongoing operational requirements.

Figure 2-3 AI Management System

- **Strategy & Policy:** Aligned with MaaS business objectives, technical capabilities and risk appetite, define the goals, scope of application, usage boundaries and risk requirements for model development, onboarding, deployment and adoption, with ongoing updates in response to technological and regulatory changes.
- **Organization, Process & Capability:** Driven by management, cross-functional teams including business, product, R&D, security, legal, privacy, compliance and operations collaborate to embed security requirements into requirement definition, R&D, testing, release, operation and decommissioning workflows, with dedicated teams, tools and expert capabilities in place.
- **Compliance & Risk Control:** Continuously identify cybersecurity, data security, personal information protection, model governance and industry regulatory requirements applicable to MaaS, and implement classification-based tiered management according to data sensitivity, model origin, application scope and potential impacts.
- **Incident Management:** Establish an incident management mechanism covering detection, alerting, analysis, response, recovery, notification and post-incident review. Beyond conventional network and data-related incidents, special attention is paid to MaaS-specific risks such as model poisoning, anomalous outputs, prompt injection, privileged-access abuse and content hazards.
- **Assurance & Continuous Improvement:** Validate the effectiveness of control measures through risk assessment, security testing, red-team exercises, internal inspections, audits and management reviews, and drive continuous improvement based on vulnerabilities, security incidents, attack samples, client feedback and regulatory updates.
- **Awareness & Culture:** Deliver role-specific training for managers, product and R&D staff, platform operators and business end-users to clarify their accountabilities for data, models, system configuration, content management and incident reporting.

2.3 AI Full-Lifecycle Security Assurance

Through the DevSecOps process, HUAWEI CLOUD embeds security requirements into the stages of data development, model development, application development, testing and validation, as well as operation and monitoring for MaaS services. This approach shifts

security left, builds-in security by design and delivers traceability across the entire chain, ensuring controllable data, trustworthy models and secure applications.

2.3.1 AI Security Pipeline Capabilities

Security capabilities are integrated into three internal R&D pipelines for data, models and applications. Risk identification and control are shifted to earlier R&D phases to ensure secure and controllable development workflows.

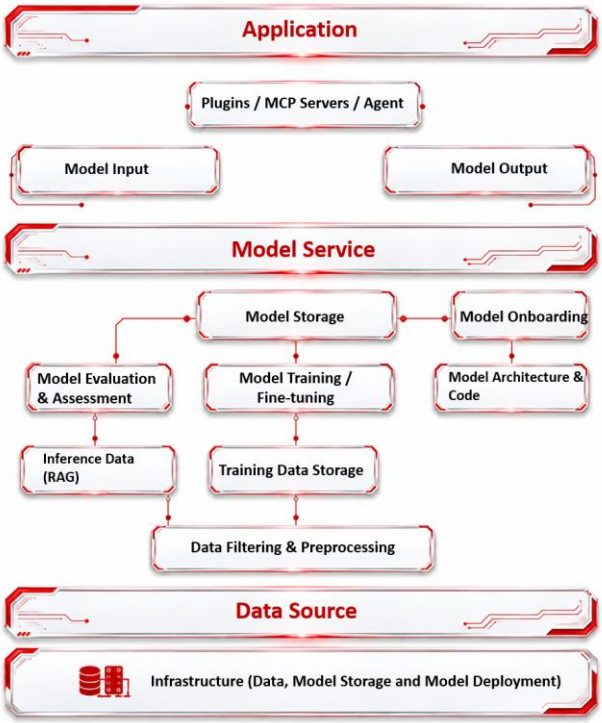

Figure 2-4 AI Application Ful Lifecycle

HUAWEI CLOUD embeds security requirements into three R&D pipelines for data, models and applications. During data development, for training data, fine-tuning data, Prompts, knowledge bases and inference input/output, measures including encryption, attack prevention, access control, compliance scanning and sensitive-information scanning are adopted to safeguard data confidentiality, integrity and compliance. In the model development phase, against risks such as data poisoning and model leakage, threat prediction, asset protection, security verification, runtime control and audit-traceability mechanisms ensure trustworthy and controllable model development. For application development, addressing the risk of uncontrolled behaviors as AI applications evolve from Q&A to Agent-based workflows, capabilities such as security isolation, access control, behavior detection and response improve the observability and governance of application behaviors.

2.3.2 Continuous Security Assurance

In addition to embedding security capabilities into the data, model and application R&D pipelines, HUAWEI CLOUD has established a security assurance mechanism covering supply-chain management, testing and validation, as well as continuous operational monitoring. Security controls are extended from the R&D phase to the full lifecycle including third-party component onboarding, product release and sustained operation, forming a closed-loop security workflow from risk identification and validation to monitoring and remediation.

Supply-Chain Security: HUAWEI CLOUD conducts automated assessment and manual review for imported third-party models, SDKs, data, frameworks and plugins. Detected security risks are promptly addressed through secure versions, patches or alternative solutions to mitigate threats introduced by third-party components and external dependencies.

Testing and Validation: HUAWEI CLOUD MaaS products undergo pre-launch testing in accordance with the *HUAWEI CLOUD Security Functional Specifications*. Regular red-team versus blue-team exercises are carried out to continuously identify security risks. Unified inspection and remediation are performed for common vulnerabilities to verify the effectiveness of relevant security controls.

Operation and Monitoring: After product launch, HUAWEI CLOUD's dedicated security operations and maintenance team adheres to the objectives of "resilience against attacks, data integrity and regulatory compliance". Supported by the large-model content risk control situational awareness platform and Security Cloud Brain, the team provides 7×24-hour security monitoring and emergency response to detect and resolve security and compliance incidents in a timely manner during runtime.

With the above mechanisms, HUAWEI CLOUD builds a continuous closed-loop security assurance covering supply-chain onboarding, pre-launch validation, runtime monitoring and emergency response. This ensures that MaaS security capabilities keep pace with product evolution, emerging risks and advancing attack techniques.

3 MaaS Shared Security Responsibility

The MaaS Shared Security Responsibility model implements the AI shared-security-responsibility principle specifically for the model-as-a-service scenario. The responsibility boundary between HUAWEI CLOUD and clients is defined based on the actual control, visibility and operability exercised by each party over cloud infrastructure, the MaaS platform, models, data and applications. HUAWEI CLOUD is accountable for "the security of the cloud and managed MaaS services themselves", while clients are accountable for "the security of content they configure, upload, develop and consume on the MaaS platform". HUAWEI CLOUD's AI shared-security-responsibility model is available for reference under **Trust Center – Responsible AI**.

Figure 3-1 MaaS Shared Security Responsibility

Security Domain	HUAWEI CLOUD Responsibility	Client Responsibility
Computing Infrastructure Security	• Responsible for chip and hardware security • Responsible for physical and network security of underlying cloud resources to ensure a trustworthy, isolated and highly available computing environment	_

Security Domain	HUAWEI CLOUD Responsibility	Client Responsibility
Model Security	• Ensure the security of environments for model deployment and inference • Provide capabilities for model evaluation, integrity protection and runtime defense • Implement security hardening for model	• Evaluate model applicability and configure security policies appropriately; bear security risks and inherent logical defects of open-source models • Ensure the legitimate and compliant use of model applications as well as content and data • Understand and abide by the license agreements of selected open-source models • Deploy additional protective measures to mitigate vulnerabilities of open-source models
Data Security	• Offer encrypted transmission channels and basic storage encryption capabilities to secure transmission links and underlying storage	The client bears primary responsibility for data security, including: • Taking charge of data provenance, authorization and compliant usage • Performing security filtering for input and output data • Implementing data desensitization and full-lifecycle encryption
Operation and Maintenance Security	• Conduct secure operations for the MaaS platform and components within its control scope, including vulnerability management, security monitoring, intrusion detection, logging, alert response, incident handling and continuous improvement; • Guarantee platform service continuity and deliver platform-side audit and traceability capabilities.	• Sustain monitoring over client-side account anomalies, model invocations, content risks and business-security incidents; configure logs, alerts and security policies as required; • Timely handle client-side incidents such as privilege abuse, abnormal invocations, content risks and application-security issues, and cooperate with HUAWEI CLOUD on necessary incident investigation, recovery and remediation.
Agent Security	• Provide security components including basic authentication and rate limiting to prevent API abuse	• Manage full-lifecycle security covering the secure development, deployment and operation of Agent applications

Table 3-1 MaaS Shared Security Responsibility Statement

4 How HUAWEI CLOUD Secures MaaS

4.1 Security Risks and Challenges for HUAWEI CLOUD MaaS

In accordance with the MaaS shared-security-responsibility boundary defined in Chapter 3, HUAWEI CLOUD is accountable for the security of the cloud and managed MaaS services themselves. During the continuous operation of MaaS services, HUAWEI CLOUD shall focus on addressing risks within its control scope, including the computing infrastructure, model services, data transmission and underlying storage, platform operation and maintenance, as well as basic Agent security components. This section outlines the key security challenges faced by HUAWEI CLOUD within the above-mentioned responsibility scope, while corresponding security assurance measures will be specified in subsequent chapters.

Figure 4-1 MaaS Security Risks and Challenges

Security Risks and Challenges of the Computing Infrastructure

The MaaS computing infrastructure features high computing density, multi-tenant sharing, complex software-hardware stacks, and long-running training and inference tasks. Security

risks can propagate along the technology stack across chips, firmware, drivers, operating systems, virtualization platforms, containers, AI frameworks and scheduling systems. Vulnerabilities, backdoors or untrusted components within chips, firmware and software modules may lead to compromised compute nodes, privilege escalation or persistent malicious code. In multi-tenant shared environments, failures in compute, memory, storage or network isolation may result in cross-tenant access, data leakage, resource interference or side-channel attacks.

Meanwhile, model weights, checkpoint files and intermediate inference results are temporarily loaded, transmitted and processed within the computing environment. Breaches of the underlying infrastructure may cause leakage, tampering or unauthorized duplication of such assets. Compute hijacking, malicious tasks, resource contention and denial-of-service attacks may degrade training and inference performance and disrupt service continuity. As GPU/NPU clusters, high-speed interconnections, distributed storage and cross-region scheduling scale up, the attack surface and dependency chains of the infrastructure expand accordingly, raising higher requirements for transmission security, trust verification, resource isolation, centralized security management and high availability. Accordingly, HUAWEI CLOUD continuously builds a trustworthy, isolated and highly available MaaS computing infrastructure at three layers: chips, hardware and AI infrastructure.

Model Security Risks and Challenges

HUAWEI CLOUD MaaS implements source verification, security evaluation and access control for pre-built third-party models, and safeguards their security during hosting, deployment and runtime on the HUAWEI CLOUD MaaS platform. Model-related risks run through the full lifecycle including development and onboarding, deployment and release, inference execution, version iteration and decommissioning. At the development and onboarding phase, risks such as training-data poisoning, security flaws in third-party models and dependent components, and malicious model files may compromise the inherent security of models. During deployment and runtime, tampering with model weights or service environments, abnormal shifts in model capability boundaries, as well as emerging attacks including prompt injection, jailbreaking and sensitive-data leakage, may undermine the security and stability of model services. Continuous version upgrades and newly exposed security vulnerabilities may introduce additional runtime risks. Therefore, HUAWEI CLOUD maintains the security of models, deployment and inference environments, and mitigates model-related risks within its control scope via model evaluation, integrity protection, runtime defense and security hardening.

Data Security Risks and Challenges

MaaS services support continuous transmission and processing of client data across platform access endpoints, inference nodes and underlying storage services. Data forms also include model context, cache, temporary files and platform-side replicas. For data-processing environments under HUAWEI CLOUD’s control, inadequate protection for transmission links, failed underlying-storage safeguards, defective tenant isolation or compromised platform access controls may expose client data to eavesdropping, tampering, unauthorized access or leakage during transmission and storage. HUAWEI CLOUD therefore prioritizes securing MaaS data transmission links and underlying storage environments to deliver baseline security guarantees for client data processed on the platform.

Security Challenges for Operations and Maintenance

Once MaaS enters persistent operation, platform-side security risks evolve dynamically and exhibit interdependency. Assets including compute resources, model versions, basic Agent services, knowledge bases, plugins and APIs are continuously created, updated and retired. Complex invocation and dependency relationships between assets raise barriers to unified governance. Meanwhile, risks such as vulnerabilities in underlying components, novel

model-targeted attacks, abnormal invocations and platform configuration changes keep evolving; one-off pre-launch assessments cannot cover emerging security issues throughout runtime.

MaaS business invocations may span models, data, basic Agent components and other platform services. Decentralized logs and intricate call chains increase the difficulty of security-event localization, impact analysis and accountability tracing. Evolving business patterns, model iterations and compliance mandates also require HUAWEI CLOUD to continuously validate the effectiveness of platform security controls. To address these challenges, HUAWEI CLOUD delivers ongoing AI asset management, risk monitoring and vulnerability management, security logging and incident handling, as well as security compliance and continuous improvement, ensuring stable and secure operation of the MaaS platform and its managed components.

Agent Security Challenges

Agent-related security risks emerge as clients build Agent applications on top of MaaS. The basic Agent interfaces and security components provided by HUAWEI CLOUD must defend against platform-side threats including authentication failures, abnormal invocations and API abuse.

4.2 Computing Infrastructure Security

4.2.1 Chip Security

Chip security serves as the foundation for the trustworthiness of the HUAWEI CLOUD MaaS computing infrastructure. Leveraging the hardware-based security capabilities of Ascend AI chips, HUAWEI CLOUD delivers trusted boot, execution isolation, data protection and trust verification for training and inference workloads.

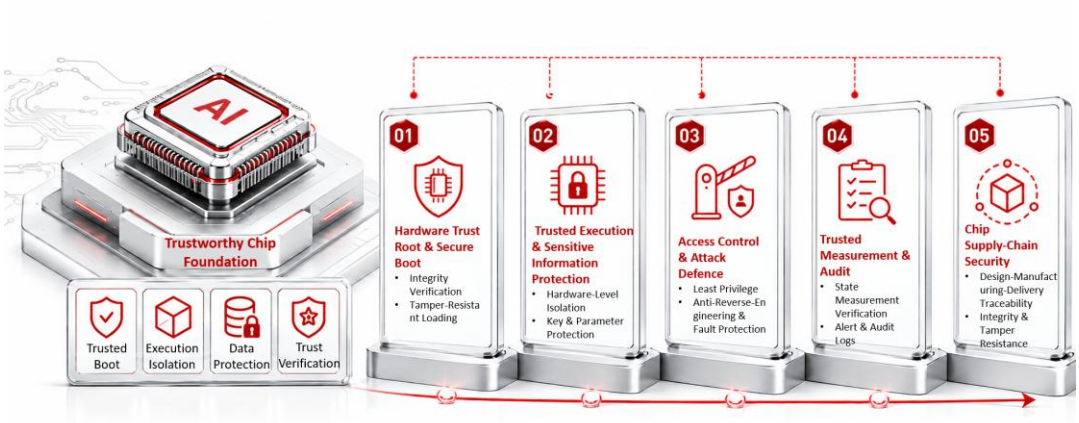

Figure 4-2 Chip Security

Hardware-based Trust Root and Secure Boot

The hardware trust root verifies the integrity of bootloaders, firmware and critical components to mitigate the risk of tampered low-level components or unauthorized code loading, establishing a trusted foundation for upper-layer servers and computing environments.

Trusted Execution and Sensitive-Data Protection

Hardware-level isolation and secure storage capabilities safeguard training data, model parameters, keys and intermediate inference results, reducing the risk of sensitive assets being accessed by host machines, other tenants or unauthorized processes.

Access Control and Attack Mitigation

Inside the chip, security domain partitioning, least-privilege access and hardware-isolation mechanisms govern access between different components and computing tasks. Combined with countermeasures against side-channel and fault-injection attacks, these features enhance the security of the runtime environment.

Trusted Measurement and Auditing

Critical hardware and firmware states are measured and validated. Alerts or blocking actions are triggered upon anomalies, while necessary measurement records are retained to support trust verification and security auditing.

Chip Supply-Chain Security

Huawei enforces security reviews, integrity checks and asset traceability throughout chip design, manufacturing, delivery and deployment. This controls risks including leakage of design assets, tampering during production and the introduction of counterfeit chips.

Security capabilities at the chip layer are built and maintained by HUAWEI CLOUD together with dedicated chip and hardware teams. Clients consume these capabilities via HUAWEI CLOUD MaaS without the need to directly manage underlying chip-level security configurations.

4.2.2 Hardware Security

Building on chip-level security, HUAWEI CLOUD leverages the security of servers, storage, network devices and physical environments to deliver trusted devices, resource isolation and service continuity guarantees for MaaS training and inference workloads.

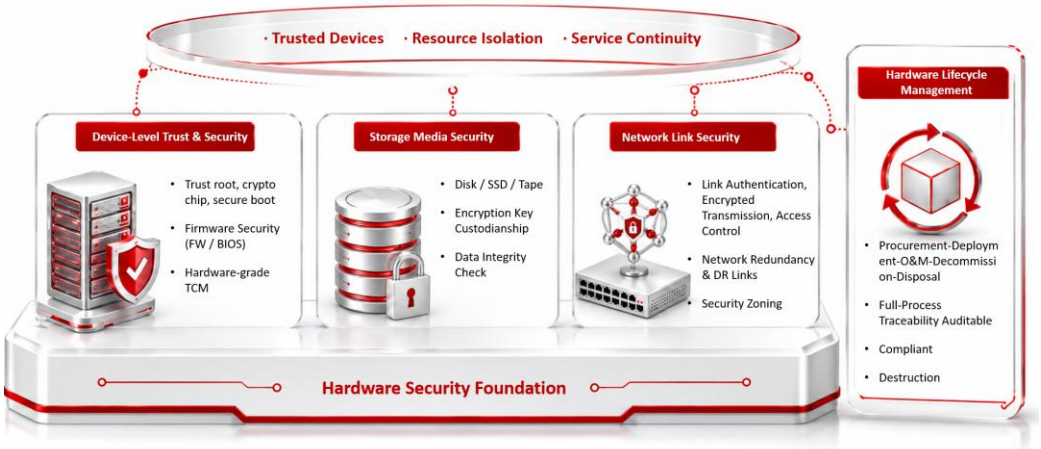

Figure 4-3 Hardware Security

Server Hardware Security

Servers establish a trust chain spanning firmware through the operating system based on the trust root, trusted boot and integrity measurement. Access control and vulnerability

management are enforced for BMC, firmware, drivers and management interfaces. Redundancy is adopted for key components to mitigate the impact of single-point hardware failures on MaaS services.

Storage Hardware Security

For training data, model files, checkpoints and client-owned data actively saved to storage resources under client control, corresponding storage services deliver encryption, integrity verification, access isolation, backup-and-recovery and quality-of-service controls to reduce risks of data leakage, tampering, loss and resource contention. Native MaaS model inference does not persist client prompts or inference outputs; such transient data is deleted immediately once inference completes.

Network Hardware Security

Identity authentication, configuration baselines, firmware upgrades and administrative access controls are applied to network devices. Network isolation, encrypted transmission, anomalous traffic monitoring and DDoS protection secure traffic for training, inference and platform management.

Hardware Lifecycle Management

End-to-end governance covers hardware procurement, acceptance, deployment, operation and maintenance, firmware and driver updates, decommissioning and media sanitization. Asset registration and status monitoring sustain a continuously trusted hardware environment.

HUAWEI CLOUD secures servers, storage, network hardware and physical data centers within its operated environment. Under the managed MaaS model, clients are not required to maintain underlying hardware directly, yet they shall configure availability zones, backup and recovery strategies appropriately in line with business continuity requirements.

4.2.3 AI Infrastructure Security

Built upon chip-level and hardware security, AI infrastructure security establishes an overall security architecture at the system layer, covering computing platforms, AI development platforms and infrastructure services.

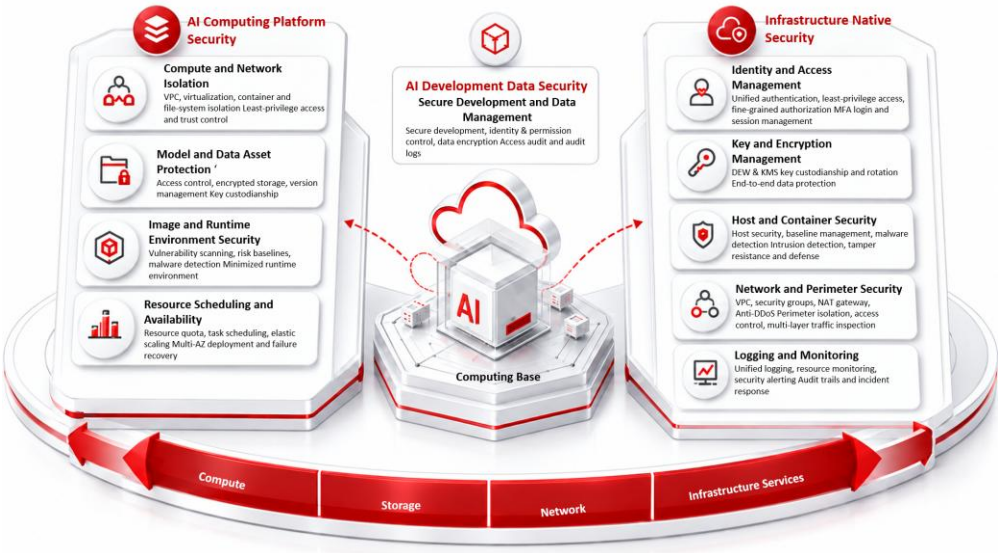

Figure 4-4 AI Infrastructure Security

4.2.3.1AI Computing Platform Security

HUAWEI CLOUD leverages AI development and computing platforms such as ModelArts to host model development, training and deployment workloads. Tenant isolation and asset protection capabilities are built across compute, network, storage, image and other dimensions.

Compute and Network Isolation

VPC, virtualization, container and file-system technologies isolate computing environments for different tenants and individual tasks. Cross-instance, cross-container and cross-network access is restricted, while runtime controls are enforced over container system calls and file access operations.

Model and Data Asset Protection

Model files, training data and intermediate artifacts can be stored in cloud services including OBS. Encrypted storage, access policies, version management and backup-and-recovery functions mitigate risks of unauthorized access, accidental deletion and data corruption.

Image and Runtime Environment Security

Vulnerability scanning, integrity verification, source validation and repository access control are performed on training and inference images to prevent images containing vulnerabilities or malicious code from entering the MaaS runtime. Vulnerability scanning, integrity verification, source validation and repository access control are performed on training and inference images to prevent images containing vulnerabilities or malicious code from entering the MaaS runtime.

Resource and Service Availability

Resource quotas, task scheduling, abnormal job monitoring, elastic scaling and failover mechanisms reduce the impact of resource contention, malicious resource exhaustion and single-point failures on training and inference services.

4.2.3.2AI Development Framework Security

AI development frameworks and their dependent components serve as a critical software layer connecting computing environments with model-development activities. HUAWEI CLOUD implements secure development, vulnerability management, version control and release integrity verification for MindSpore and associated frameworks, dependency libraries and components. These measures mitigate risks whereby framework vulnerabilities, malicious dependencies and untrusted components compromise the MaaS training and inference environments.

Model-level capabilities including adversarial training, model robustness, privacy-preserving training and model explainability are elaborated in Chapter 4 (Model Security). Relevant authoritative certifications are listed collectively in Chapter 9.

4.2.3.3Infrastructure Service Security

HUAWEI CLOUD delivers foundational security capabilities covering identity, keys, hosts, containers and networks for the MaaS computing environment via cloud-native security services:

- Identity and Access Control: IAM provides unified authentication, multi-factor authentication, fine-grained authorization, access keys and identity federation to govern user and service access to MaaS resources.

- Key and Encryption Management: DEW, KMS and hardware security modules protect encryption keys to support encryption for model files, training data and cloud storage.
- Host and Container Security: Capabilities such as HSS deliver asset discovery, vulnerability management, baseline inspection, intrusion detection and container runtime protection.
- Network and Perimeter Security: VPC, security groups, network ACLs and Anti-DDoS enable network isolation, access control, anomalous traffic detection and protection against denial-of-service attacks.
- Logging and Monitoring: Platform access, resource changes and security alerts are logged and fed into a unified security operations platform for continuous monitoring and incident response.

Clients shall configure account permissions, VPC and security group policies, access keys, images and dependent components according to their business requirements, and assume responsibility for the security of their computing tasks and custom runtime environments. HUAWEI CLOUD ensures the sustained operation of the managed platform, underlying services and their built-in security mechanisms.

4.3 Model Service Security

HUAWEI CLOUD MaaS delivers platform capabilities for model development, onboarding, deployment and inference for Huawei-developed models, pre-built third-party models and client-customized models. In accordance with the MaaS shared-security-responsibility boundary defined in Chapter 3, HUAWEI CLOUD prioritizes securing models, deployment and inference environments within its control scope, and provides capabilities including model evaluation, integrity protection, runtime defense, model security hardening and vulnerability remediation.

Targeting the model-security objectives of **controllability, compliance and evaluability**, HUAWEI CLOUD embeds mechanisms covering R&D governance, supply-chain validation, model asset protection, security assessment, release control and runtime protection across three phases: model development and onboarding, deployment and release, as well as runtime. Differentiated security safeguards are implemented based on model origins and service modes.

4.3.1 Security for Model Development and Onboarding

HUAWEI CLOUD MaaS adopts differentiated security management based on model sources. For self-developed HUAWEI CLOUD models, security requirements run through the entire lifecycle including data engineering, model training, security alignment, model evaluation and version release. For pre-built third-party models, HUAWEI CLOUD performs source verification, supply-chain inspection, security assessment and platform admission before offering them as MaaS model services. For client-customized models, HUAWEI CLOUD mitigates security risks posed to the MaaS managed environment mainly through integrity checks, security evaluation and platform admission controls. Models imported into the production-ready MaaS model asset repository are assigned dedicated model identifiers, version tags, security status labels and associated records to support subsequent deployment, runtime operations and incident traceability.

Subject to model origin, controllable scope and practical service capabilities, HUAWEI CLOUD conducts platform admission assessments and risk evaluations for pre-built third-party models, and provides clients with available models or service descriptions. Evaluation results reflect risk profiles only for specific versions, testing scopes and test conditions, and shall not be construed as a warranty for the security, accuracy, authenticity, completeness or business fitness of third-party models and their outputs.

4.3.1.1Model Development Security

HUAWEI CLOUD’s self-developed models are represented by the Pangu Large Model. The Pangu Large Model adopts a layered architecture consisting of L0 base models, L1 industry-specific models and L2 scenario-oriented models. The L0 tier covers five categories of base models: natural language, computer vision, multimodal, prediction and scientific computing. HUAWEI CLOUD embeds security requirements into the full lifecycle of its self-developed models, including data preparation, model training, security alignment, model evaluation, version release and defect response, while retaining traceable links among datasets, training jobs, model weights, evaluation results and released versions.

ModelArts Studio provides data engineering and model-development capabilities for Pangu Large Model research and development. Data engineering spans data acquisition, processing, synthesis, labelling, proportioning, evaluation and release. It generates process records covering the transformation from raw data to training datasets, delivering platform-enabled support for data quality assurance and traceable governance throughout model development.

Data Governance and Labelling

For HUAWEI CLOUD self-developed models, R&D data governance and labelling mainly cover the following requirements:

- The source, license, intended usage, sensitivity level, retention period and decommissioning criteria shall be documented for training and evaluation data. Data with unknown provenance, insufficient authorization or mismatched usage purposes shall not be admitted into the production-grade R&D pipeline.
- Data quality checks, identification of duplicate and anomalous samples, detection of contamination and potential poisoning, as well as manual reviews are performed to mitigate the risk of poor-quality or malicious data skewing model behavior.
- Least-privilege access is enforced for labelling accounts by project and role. Sensitive data is de-identified on a task-specific basis. A closed-loop quality assurance mechanism is implemented through multi-annotator labelling, result reviews, consistency checks and audits of anomalous annotations.

【**HUAWEI CLOUD Service**】**ModelArts Studio** Data Engineering supports data acquisition, processing, synthesis, labelling, proportioning, evaluation and release. Data labelling covers labelling tasks, review tasks and task management. Depending on data types, it provides multi-annotator labelling, annotation review, task handover and AI-powered pre-labelling for text content.

Environment, Supply-Chain and Weight Security

- **R&D Environment Isolation:** HUAWEI CLOUD segregates R&D, test and production environments for its self-developed model development and model-processing environments under MaaS control. Boundaries for projects, networks, computing, storage, identities and keys mitigate risks of unauthorized cross-environment access and credential abuse.
- **Model Supply-Chain Security:** For third-party libraries, frameworks, images and components introduced during the development of HUAWEI CLOUD's self-built models, as well as third-party and client-customized models to be deployed in the MaaS managed environment, necessary security checks are performed against source information, integrity, malicious files, vulnerabilities, custom operators and executable codes based on object types. This reduces risks posed by untrusted components to MaaS model development and runtime environments.
- **Model Asset Protection:** Access control, transmission and storage protection, integrity verification and audit trails for critical operations are enforced for model weights, checkpoints and adapters managed and hosted by HUAWEI CLOUD. These measures mitigate the risks of unauthorized access, tampering or leakage of model assets throughout development, deployment and runtime.

【**HUAWEI CLOUD Service**】**Identity and Access Management (IAM)** enables isolation for users, user groups, agencies and projects, as well as least-privilege control.

【**HUAWEI CLOUD Service**】**Data Encryption Workshop (DEW)** delivers key custodianship, encryption and rotation capabilities to mitigate unauthorized reading of weights, datasets and associated credentials.

【**HUAWEI CLOUD Service**】**Cloud Trace Service (CTS)** logs administrative operations on core cloud resources to support traceability for changes to model assets and permissions.

4.3.1.2Model Screening and Onboarding Process

For models intended to be offered as pre-built third-party models or uploaded by clients to the MaaS managed environment, HUAWEI CLOUD performs corresponding security admission controls based on the model source type and platform governance scope. The workflow covers candidate model registration, source verification, integrity and supply-chain inspection, security evaluation, as well as risk classification and admission approval. For risks that may compromise the security of the MaaS platform or managed services, countermeasures including admission rejection, mandatory remediation or service-scope restrictions shall be adopted according to risk levels.

Self-developed HUAWEI CLOUD models complete security admission via internal R&D, evaluation and release procedures. Pre-built third-party models go through platform-specific onboarding and admission governance administered by HUAWEI CLOUD. For client-customized models, HUAWEI CLOUD focuses on platform-side security admission, while clients remain accountable for model provenance, licensing and business fitness in line with the responsibility boundaries defined in Chapter 3.

The workflow described below primarily applies to the onboarding of pre-built third-party models and client-customized models. Self-developed HUAWEI CLOUD models obtain security clearance through internal R&D, evaluation and release processes.

Figure 4-5 Model Screening and Onboarding Process

The table below describes the key control activities for model onboarding and security admission on the HUAWEI CLOUD MaaS platform. The detailed responsibilities shall be subject to the responsibility boundaries specified in Chapter 3.

Process Step	Key Checks	Output
Candidate Model Registration	Model name, version, source, intended use, responsible owner, target users and data scope	Candidate model record
Source and License Review	Provider identity, license terms, restrictions on commercial use and redistribution, geographic and industry-specific constraints	License review conclusion
Integrity and Supply-Chain Inspection	Signature / hash value, model file format, vulnerabilities in images and dependencies, malicious code, custom operators	Scan report and bill of materials

Process Step	Key Checks	Output
Model Security Evaluation	Content safety, privacy reproduction, robustness, factuality, bias and fairness, jailbreak resistance and model extraction resistance	Evaluation report and defect list
Risk Classification and Approval	Business impact, user coverage, data sensitivity, capability boundaries, compensatory controls and residual risks	Admission, remediation, restriction or rejection decision
Model Asset Warehousing	Associate models, weights, images, evaluation records, model cards, responsible owners and version identifiers	Traceable model assets

Figure 4-1 Model Screening and Onboarding Model Screening and Onboarding statement

4.3.2 Security for Model Deployment and Release

For MaaS model services delivered or hosted by HUAWEI CLOUD, HUAWEI CLOUD establishes mechanisms for security admission, risk assessment, release approval, version management, canary validation, rollback and decommissioning prior to model go-live. Release artifacts are associated with model weights, inference environments, security policies, evaluation results and version identifiers to ensure the model deployment and release lifecycle is identifiable, verifiable and traceable, with rollback or decommissioning capabilities available in the event of critical security incidents.

4.3.2.1Security Admission and Evaluation

- HUAWEI CLOUD conducts targeted security evaluations for candidate or managed MaaS model services based on model type, service scope and risk profile, and defines mandatory release gates, defect remediation rules and residual risk management requirements. Evaluations may cover dimensions including content safety, privacy protection, robustness, factuality, bias and fairness, and model capability boundaries.
- Red-team testing covers jailbreaking, direct and indirect prompt injection, system prompt leakage, sensitive-information reproduction, harmful outputs, model extraction and multimodal bypass attacks. Records are maintained for attack library versions, test coverage, defect severity ratings and retesting exit criteria.
- Dedicated testing is performed for financial, healthcare, government, education and minor-facing scenarios. For model services deployed in high-risk or heavily-regulated domains, specialized assessments and risk validation are carried out in line with the actual service scope.
- Model descriptions or service specifications are prepared for HUAWEI CLOUD-hosted public model services. Model versions, core capabilities, applicable scenarios, known limitations and mandatory risk disclosures are provided according to real-world service conditions to deliver baseline information for client awareness and model selection.

4.3.2.2HUAWEI CLOUD Business Practices

HUAWEI CLOUD MaaS implements differentiated management based on model sources. Self-developed models, pre-built third-party models and client-customized models are all governed under unified model asset, evaluation and release management frameworks, while responsibility boundaries, inspection depth and available open capabilities are defined according to model origins and risk levels.

【HUAWEI CLOUD Service】MaaS hosts model assets, online inference and relevant platform management capabilities, delivering corresponding platform support for model

deployment, release and runtime operations. The exact scope of support is subject to the target region, management console and official product documentation.

【**HUAWEI CLOUD Service**】**Cloud Trace Service (CTS)** logs critical administrative operations for MaaS and associated cloud resources to enable traceability for releases and configuration changes.

4.3.2.3Release, Canary Deployment and Rollback

- After approval for model release, low-traffic canary testing or shadow evaluation is adopted first to verify security alert hits, false interception, factuality, response latency and failure rate. Traffic scale-up is permitted only after exit criteria are satisfied.
- A new security evaluation shall be triggered upon material changes to model weights, inference images, system prompts or content security policies. Simplified procedures may apply for low-risk cosmetic adjustments.
- Validated versions, configurations and rollback paths shall be retained. Capabilities for traffic switching, legacy version blocking and emergency decommissioning are verified periodically.

4.3.3 Model Runtime and Application Security

Once a model enters the online inference and continuous service phase, HUAWEI CLOUD prioritizes securing models, inference environments and MaaS model services within its control scope. Leveraging native model-security capabilities, inference-chain security mechanisms, platform-based identity and access control, tenant-resource isolation, and ongoing remediation of model-security flaws, HUAWEI CLOUD mitigates runtime risks including prompt injection attacks, anomalous outputs, sensitive-information disclosure, unauthorized access and model-related vulnerabilities.。

Across the end-to-end MaaS model application pipeline, HUAWEI CLOUD also provides configurable safety guardrails and associated cloud-security capabilities for client selection. Business permissions, security policies and end-usage requirements on the client side are elaborated in Section 5.2.

The diagram below illustrates the multi-layer protection framework covering the full lifecycle from model input and inference through to output and application execution. HUAWEI CLOUD is accountable for platform and model runtime security controls within its governance boundary. Client-application-side controls shown in the figure shall be implemented by clients in accordance with the responsibility boundaries defined in Chapter 3, with further details provided in Section 5.2.

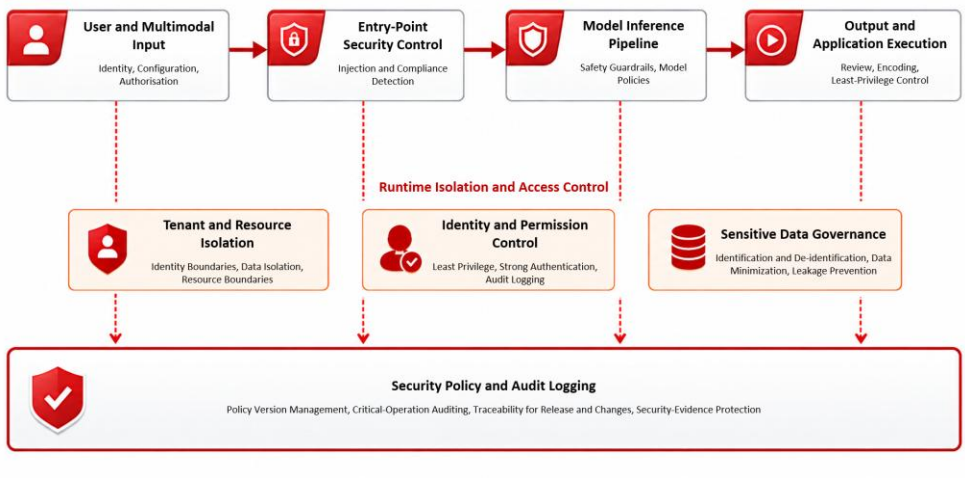

Figure 4-6 Layered Protection for Model Runtime and Application Security

4.3.3.1Content Security

Within the model inference pipeline, Huawei Cloud MaaS provides Huawei Cloud LLM Firewall and corresponding input-output security capabilities. It performs security detection on prompts, context, and images fed into the model, as well as texts, images and other content generated by the model. This mitigates impacts on MaaS services and client businesses caused by malicious inputs, non-compliant content, sensitive-information leakage and anomalous model outputs. Relevant capabilities can be configured according to model types, application scenarios, deployment regions and actual service availability.

Figure 4-7 Large Model Firewall

In terms of text content detection, the HUAWEI CLOUD Large Model Firewall identifies pornographic, violent-terrorist, prohibited, malicious and other non-compliant information within inputs and outputs. It also monitors exposure risks for personal identifiers, device data, account credentials and enterprise-sensitive information. For multilingual scenarios, targeted

detection and response are performed within the supported language scope to mitigate bypass attempts leveraging minor-language or mixed-language payloads.

For model-specific semantic attacks, built-in security capabilities detect risks such as direct prompt injection, jailbreak attempts, system-prompt probing, reverse prompting, malicious instructions and injected code. In RAG, Memory and Agent use cases, the solution addresses indirect prompt injection concealed within external documents, knowledge-base entries and persistent memory. This prevents untrusted content from skewing model behavior, disclosing sensitive data, evading predefined guardrails or triggering unintended tool invocations. For Agents integrated with tools and business systems, dedicated Agent-security controls detect malicious prompting, tool abuse and Agent hijacking. Refer to Section 4.6 for detailed safeguards.

For image and multimodal content inspection, the HUAWEI CLOUD Large Model Firewall analyses image semantics and embedded OCR text within the scope of service capabilities. It detects policy-violating material, sensitive data and hidden instructions carried in visuals, while guarding against cross-modal bypass risks constructed from combined text-image payloads. Comprehensive multi-modal evaluation reduces the likelihood of attackers evading pure-text detection via embedded text, visual cues or modality-specific evasion techniques.

Within HUAWEI CLOUD MaaS, independent input and output inspection is supported subject to service specifications, alongside continuous scanning for streaming responses. This minimizes the exposure window associated with post-generation-only validation. Capabilities can be plugged into the model invocation chain via APIs, forming a multi-layered defense against conventional web exploits, malicious requests and model-specific semantic threats when combined with entry-point controls including WAF and API Gateway.

HUAWEI CLOUD enables business-oriented content-security policy configuration. Response actions including blocking, logging, desensitization and safe fallback replies can be defined for detection categories such as prohibited content, prompt injection and sensitive-data leakage. Distinct policy bundles may be assigned across applications, models and risk profiles. Exemption rules and policy-tuning workflows reduce false-positive disruption to legitimate workflows. Clients shall configure detection scopes, enforcement actions and business-side human-review workflows in alignment with their use cases, user demographics and compliance obligations.

Content-security inspections generate structured security logs and risk records that capture threat categories, inspection dimensions and disposition outcomes, with aggregated statistics for different attack vectors and content risks. Based on the platform’s observable range, HUAWEI CLOUD monitors inspection volumes, block counts and risk trends to support security alerting, incident investigation and iterative policy refinement.

【HUAWEI CLOUD Service】The HUAWEI CLOUD **Large Model Firewall** provides security detection capabilities for text, images and other content types to identify non-compliant material and associated content risks. The actual supported detection categories, languages, input/output modes, response actions and regional coverage are subject to the target region, management console and official product documentation.

4.3.3.2Runtime Isolation and Access Control

On the MaaS platform side, HUAWEI CLOUD enforces necessary isolation for model assets, data and inference resources belonging to different tenants. Platform-enabled identity authentication, model-service access control and resource-isolation mechanisms mitigate risks such as cross-tenant access, unauthorized invocation and mutual resource interference. Service-access protection is applied to model hosting, online inference and associated management interfaces to secure model runtime environments and service interfaces within HUAWEI CLOUD’s control scope.

Security administration for configurable items including client accounts, roles, API Keys, network access policies, traffic and quotas shall be set up by clients in accordance with their business requirements. HUAWEI CLOUD provides supporting platform and cloud-service capabilities, which are detailed in Section 5.2.

4.3.3.3Model Runtime Security Boundaries

- For model services provided or hosted by HUAWEI CLOUD, in the event of material changes to model capability boundaries, critical security flaws or major risks that may compromise the secure operation of MaaS, HUAWEI CLOUD shall adopt countermeasures including security hardening, protection-policy adjustment, capability restriction, rollback to validated versions or decommissioning of the relevant model services based on risk levels. Model-related security issues are tracked and remediated continuously within operational security workflows.
- Upon retirement of model services delivered and governed by HUAWEI CLOUD, corresponding service endpoints will be disabled. Model weights, adapters, runtime images, caches and other hosted assets will be disposed of in compliance with platform asset and data management requirements, mitigating risks of residual exposure or unauthorized usage for decommissioned models.

4.4 Platform Data Transmission and Storage Security

Data security serves as a horizontal security capability within the HUAWEI CLOUD AI Security Framework. From the perspective of the three data states, MaaS data security covers data-in-transit, data-at-rest and data-in-use. In accordance with the MaaS shared security responsibility boundary defined in Chapter 3, HUAWEI CLOUD focuses on securing data transmission links and underlying storage within its control scope. It protects data-in-transit and data-at-rest through capabilities such as secure transmission, network isolation, storage encryption and key protection. For data-in-use that enters computing and inference environments, HUAWEI CLOUD safeguards the processing environment via tenant-level, job-level and runtime-environment isolation built into the computing infrastructure. Clients assume corresponding responsibilities for their own data provenance, authorization, compliant usage and specific data protection policies. Further details are provided in Section 5.3.

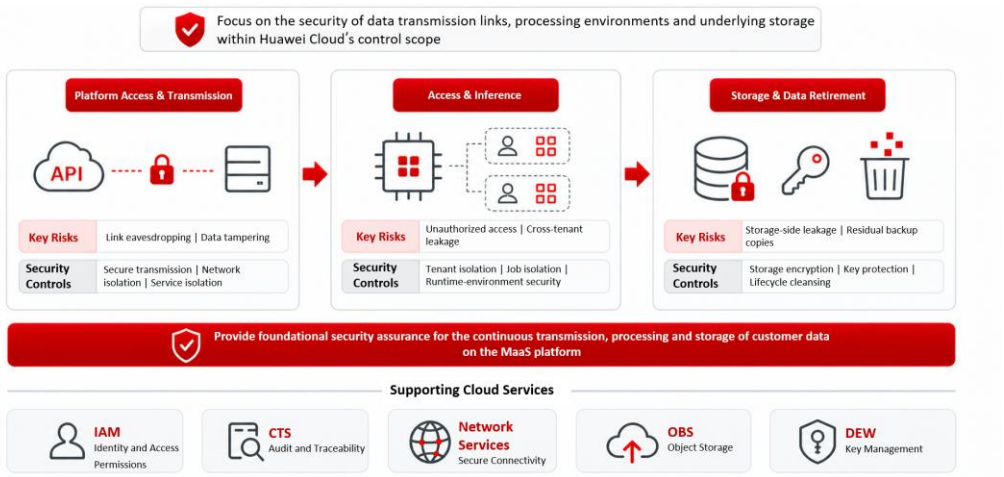

Figure 4-8 Platform Data Transmission and Storage Security

4.4.1 Data-in-Transit Security

Client data flows into HUAWEI CLOUD MaaS via APIs, object storage and other channels, and circulates across the MaaS platform, storage services, as well as associated training and inference environments. HUAWEI CLOUD enforces security protection for data transmission links within its control scope. Mechanisms including secure communication protocols, network isolation and service isolation safeguard the confidentiality and integrity of data during access and internal platform circulation.

For diverse deployment scenarios, HUAWEI CLOUD delivers secure data connectivity options. Clients may connect to MaaS services over encrypted public-network links or controlled networks based on their architecture, which reduces exposure risks during cross-network and cross-service data transmission. The scope, content and specific usage requirements of client-side data transmission are managed by clients for their business needs; further details are provided in Section 5.3.

Once data enters MaaS computing and inference environments, its state shifts from data-in-transit to data-in-use. Leveraging tenant isolation, job isolation and runtime security controls offered by the computing infrastructure and the MaaS platform, HUAWEI CLOUD protects data within memory, video memory, caches and relevant processing environments, mitigating risks of unauthorized data access across different tenants and jobs. Security capabilities of the underlying computing foundation are described in Section 4.2 and will not be repeated in this section.

【**HUAWEI CLOUD Service**】**Identity and Access Management (IAM)** delivers identity authentication and fine-grained permission management capabilities to support identity and access control for MaaS-related users and resources. The actual scope of support is subject to the target region, management console and official product documentation.

【**HUAWEI CLOUD Service**】**Cloud Trace Service (CTS)** logs critical administrative operations on relevant cloud resources, enabling audit traceability for data ingestion, permission adjustments and resource changes. The set of auditable events is defined in the official product documentation.

【**HUAWEI CLOUD Service**】HUAWEI CLOUD Network Services provide secure data connectivity for diverse scenarios via **VPC** endpoints, **VPN** or **Direct Connect**. The specific connection method is determined by the real-world deployment architecture.

4.4.2 Data Storage and Destruction Security

MaaS may generate job-specific data and hosted assets during training, evaluation, model hosting, and storage scenarios explicitly configured by clients. HUAWEI CLOUD performs cleansing and decommissioning management for such data and resources in compliance with applicable data retention policies and resource lifecycle requirements. For native MaaS model inference, client prompts and model inference outputs are processed only during the inference session without persistent platform copies, and are deleted immediately after inference completes. Data such as cache entries, temporary files and platform replicas may be generated during MaaS service operation. HUAWEI CLOUD implements retention, cleansing and decommissioning controls for data and associated resources within its governance scope in line with MaaS data retention rules and platform resource lifecycle policies, ensuring platform-side data goes through full lifecycle handling as defined.

HUAWEI CLOUD secures hosted storage and underlying storage environments under its control, and applies baseline storage encryption for data based on native platform and storage service capabilities. Clients remain responsible for data sensitivity classification, self-configurable encryption policies, retention periods and business-side data copies. Relevant requirements are specified in Section 5.3.

【HUAWEI CLOUD Service】**Object Storage Service (OBS)** provides foundational storage capabilities including object storage and server-side encryption for MaaS-related data. The actual supported scope shall be subject to the target region, management console and official product documentation.

【HUAWEI CLOUD Service】**Data Encryption Workshop (DEW)** delivers key management and encryption capabilities to support key security for protecting MaaS-related data-at-rest. The detailed scope of support is defined in the official product documentation.

4.5 Platform Sustained Security Operations

After MaaS enters the continuous-operation phase, computing resources, MaaS platform components, model services and Agent-related underlying services within HUAWEI CLOUD's control scope keep changing in terms of assets, versions, configurations and runtime states. Security risks also evolve alongside platform upgrades and emerging attack techniques. In accordance with the MaaS shared-security responsibility boundary defined in Chapter 3, HUAWEI CLOUD carries out ongoing security operations for the MaaS platform and hosted components under its governance to ensure the sustained effectiveness of relevant security controls, monitoring capabilities and platform services.

HUAWEI CLOUD builds a closed-loop operational security system covering four domains: AI asset management, risk monitoring and vulnerability management, security log and incident management, as well as security compliance and continuous optimization. It delivers the workflow of "Asset Visibility – Risk Identification – Incident Response – Continuous Improvement" to sustain the secure operation of the MaaS platform and its in-scope components.

Figure 4-9 MaaS Sustained Secure Operation

4.5.1 AI Asset Management

AI asset management serves as the foundation for HUAWEI CLOUD's sustained security operations for MaaS. As the MaaS platform undergoes iterative updates, AI assets within HUAWEI CLOUD's scope of development, operation and hosting — including computing resources, platform components, models and their versions, model services, APIs and foundational Agent services — are continuously added, modified and decommissioned, with dynamic shifts in asset status and dependency relationships. Through a unified asset management mechanism, HUAWEI CLOUD maintains ongoing visibility into the type,

version, runtime status, responsible entity, security posture and associations of such assets, reducing management blind spots caused by the rapid evolution of AI assets.

HUAWEI CLOUD establishes a centralized asset inventory for AI assets that it develops, operates and hosts. The inventory primarily covers computing power and runtime environments, MaaS platform components, models and corresponding versions provided or hosted by HUAWEI CLOUD, inference services, platform-native APIs, foundational Agent services and associated security components. It records asset type, responsible party, version, deployment status, security posture, risk level and upstream-downstream dependencies, delivering a consistent asset baseline for vulnerability management, risk monitoring, incident analysis and security policy enforcement.

HUAWEI CLOUD keeps MaaS asset records up-to-date using built-in asset discovery and identification capabilities. When assets are provisioned, version-upgraded, reconfigured, redeployed or retired, their statuses are updated synchronously. Dependency and invocation relationships among computing resources, platform components, model services, APIs and foundational Agent services are mapped to provide a continuously refreshed asset view for risk assessment and security operations.

Version and change logs are maintained for model iterations, platform components, APIs and related MaaS services. Major version upgrades, critical component adjustments, high-privilege services and other modifications that may alter security boundaries trigger security reviews in accordance with risk levels. Unregistered components, anomalous assets and ungoverned services that are detected are promptly onboarding into asset management, with reviews, access restrictions or remediation actions applied based on risk severity.

Differentiated management is implemented according to asset criticality, service privileges, security status and business impact. Priority monitoring and enhanced security controls are applied to platform assets that deliver core MaaS capabilities, expose high-privilege interfaces or carry significant business impact.

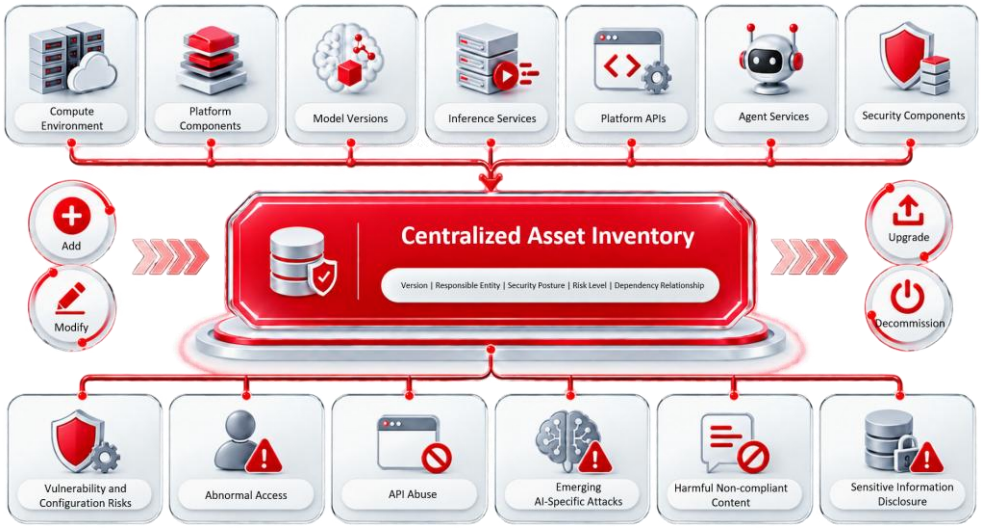

Figure 4-10 AI Asset Dynamic Management and Risk Monitoring

4.5.2 Risk Monitoring and Vulnerability Management

Risk monitoring and vulnerability management are designed to continuously detect and remediate security risks arising during the operation of the MaaS platform and hosted

components within HUAWEI CLOUD's control scope. HUAWEI CLOUD performs correlated monitoring based on platform asset information, security logs, vulnerability data, security alerts and runtime status to identify anomalous platform access, component vulnerabilities, emerging AI-specific attacks, content compliance risks and service anomalies. A closed-loop workflow is established through risk assessment, alerting, handling, remediation and validation.

HUAWEI CLOUD builds a unified risk monitoring view covering the computing infrastructure, MaaS platform, model services and foundational Agent services under its governance, with correlations mapped across assets, vulnerabilities, access behaviors, security alerts and operational states. At the computing infrastructure and platform layers, ongoing monitoring covers vulnerabilities, unauthorized access and resource anomalies for compute nodes, containers, images, networks and platform components. At the model- and Agent-enabled platform capability layer, within observable boundaries, the system detects abnormal requests, attack activities and interface abuse targeting MaaS services, alongside content-security and compliance risks such as anomalous outputs, harmful or non-compliant material and sensitive-information disclosure. Corresponding security alerts are generated to support risk localization and impact analysis.

To address complex invocation chains and scattered risk signals in MaaS, HUAWEI CLOUD correlates identity authentication records, platform access logs, model-service invocations, component runtime metrics, security-detection results and vulnerability data within its observable scope. Security signals generated by disparate components are mapped to relevant assets and services, helping identify the source, blast radius and associated risks of suspicious activities. For security incidents requiring deeper investigation, the security-log and incident-management capability performs call-chain tracing and incident response.

In terms of content compliance, clients bear primary responsibility for content governance. As the MaaS-platform provider, HUAWEI CLOUD supports clients in lawful and compliant use of content data. In accordance with applicable laws, regulatory requirements and the *MaaS-Model-as-a-Service Statement of Service*, HUAWEI CLOUD conducts necessary security and compliance reviews and response actions for client content via automated or manual means. The platform may perform real-time content-security scanning on model inference inputs and outputs against categories including politically sensitive material, pornography, violent content and fraud-related information. When risk rules are triggered, the platform records non-content event metadata within the bounds of the service statement and product capabilities — such as risk category, detection dimension, disposition outcome, model or service version, event timestamp and hit count. Raw client input-output payloads and content snippets are not retained. In cases of critical compliance risks, the platform delivers content-risk notifications to tenants through client managers, specifying violation types, trigger counts and affected models together with recommended mitigations. The platform automatically enforces strengthened content-security policies for impacted models, with temporary restrictions on high-risk generation capabilities applied where necessary.

HUAWEI CLOUD maintains continuous vulnerability discovery and risk assessments for compute nodes, containers, images, AI frameworks and associated MaaS-platform components that it operates and hosts. Remediation actions are prioritized based on vulnerability severity, exploitability, asset criticality and business impact. Validation checks are completed after fixes are deployed. For high-risk vulnerabilities that may compromise MaaS-platform security or service continuity, mitigation, patching, isolation and other countermeasures are applied in a timely manner.

HUAWEI CLOUD iteratively refines platform-side detection rules and security-protection policies using emerging attack samples, vulnerability intelligence, red-team exercises, security-incident data and platform change records. Policy validation is performed where required to minimize service disruption to legitimate MaaS workloads caused by security-policy updates.

4.5.3 Risk Monitoring and Vulnerability Management

MaaS business invocations may span identity authentication, platform interfaces, model services, data processing and foundational Agent components. Dispersed logs and complex invocation chains increase the difficulty of security-incident localization and impact analysis. Within its control and observability boundaries, HUAWEI CLOUD establishes platform-native logging, correlation analysis and security-incident response mechanisms. By correlating critical operations, security alerts and invocation data, HUAWEI CLOUD enhances the observability, traceability and response capability for security incidents on the MaaS platform.

HUAWEI CLOUD logs necessary information including identity-authentication records, resource-access trails, model and service versions, configuration changes, API invocations, security-detection results, administrative operations and security alerts in line with actual platform observability. Sensitive content within logs is properly minimized, desensitized and protected via access-control safeguards.

Upon detection of platform-side security anomalies, relevant logs can be correlated across dimensions such as tenant, user, model, service, timestamp and incident type. Analysis of critical invocation paths helps pinpoint incident sources, affected assets and blast radius, delivering platform-based evidence for incident investigation and accountability.

HUAWEI CLOUD implements a security-incident response workflow covering alerting, triage, severity classification, remediation, recovery and post-incident review. For security events within HUAWEI CLOUD's control scope — including vulnerability exploitation, anomalous access, platform-targeted attacks and service failures — appropriate risk-based countermeasures are taken, such as access restrictions, policy adjustments, component remediation, service isolation or service recovery. For incidents crossing the security boundary between HUAWEI CLOUD and clients, necessary information sharing and joint response will be performed within the respective controllable and observable scopes of both parties, in accordance with the shared-responsibility principles defined in Chapter 3.

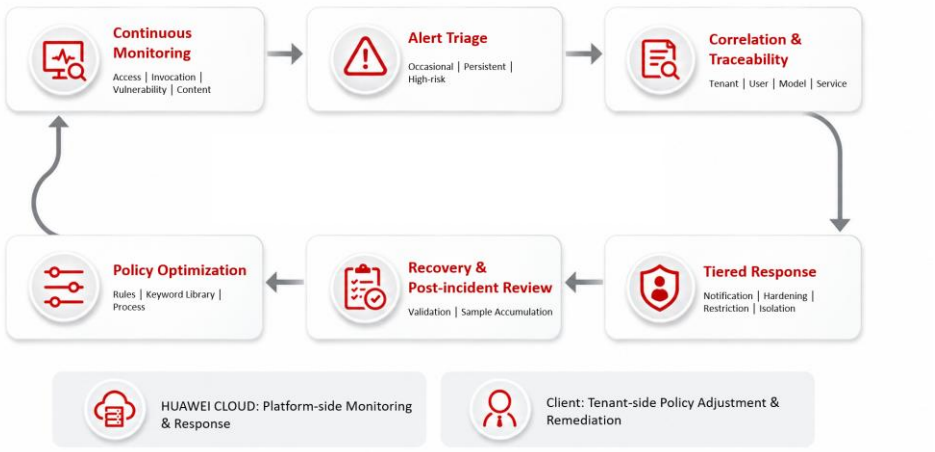

Figure 4-11 MaaS Security Event Management

4.5.4 Security Compliance and Continuous Optimization

Security compliance and continuous optimization form the closed-loop security operation for HUAWEI CLOUD MaaS. HUAWEI CLOUD keeps track of cybersecurity, data security, personal information protection, AI governance and other applicable regulatory requirements

governing its MaaS platform and services. It continuously evaluates the effectiveness of its security controls against platform assets, vulnerabilities, security incidents, attack samples and service changes.

Based on platform-level risks, major version upgrades, critical component modifications and applicable regulatory mandates, HUAWEI CLOUD conducts necessary security testing, risk assessments, red-team exercises, compliance audits and other validation activities for the MaaS platform, underlying computing infrastructure, model services and foundational Agent services. Identified issues are remediated and re-validated accordingly.

Drawing on vulnerability intelligence, attack samples, security incidents, red-team results and platform updates, HUAWEI CLOUD iteratively enhances detection rules, security policies and operational workflows, enabling security capabilities to evolve alongside the MaaS platform and the changing threat landscape.

Subject to service features and applicable scopes, HUAWEI CLOUD may provide supporting documentation including security evaluations, certifications and audit evidence to assist clients with their internal security assessments and compliance reviews. Relevant certifications and supporting materials are consolidated in Chapter 7.

4.6 Agent Platform Basic Security Measure

After clients build Agent applications on top of HUAWEI CLOUD MaaS, Agents may invoke models and associated services through platform interfaces. In accordance with the MaaS shared-security responsibility boundary defined in Chapter 3, HUAWEI CLOUD delivers platform-side foundational security components for Agent security, including identity authentication, access control, quota management and rate limiting. These controls mitigate security impacts on the MaaS platform and hosted services caused by unauthorized access, anomalous invocations and API abuse. Clients remain accountable for the security of the Agent application itself, covering identity-permission design, knowledge-base access, Skill/MCP and tool invocation, business execution and ongoing operational security. Details are provided in Section 5.5. After clients build Agent applications on top of HUAWEI CLOUD MaaS, Agents may invoke models and associated services through platform interfaces. In accordance with the MaaS shared-security responsibility boundary defined in Chapter 3, HUAWEI CLOUD delivers platform-side foundational security components for Agent security, including identity authentication, access control, quota management and rate limiting. These controls mitigate security impacts on the MaaS platform and hosted services caused by unauthorized access, anomalous invocations and API abuse. Clients remain accountable for the security of the Agent application itself, covering identity-permission design, knowledge-base access, Skill/MCP and tool invocation, business execution and ongoing operational security. Details are provided in Section 5.5.

HUAWEI CLOUD enforces mandatory identity verification for Agent access to MaaS-related platform services via platform-native identity authentication and access-control mechanisms. Quota and rate-limiting rules for platform interfaces constrain abnormal or excessive calls, reducing risks to platform security and service stability arising from credential misuse, malicious requests and excessive API resource consumption. Detected anomalous interface invocations within HUAWEI CLOUD's control scope are monitored and remediated through platform security-operation workflows.

【HUAWEI CLOUD Service】Identity and Access Management (IAM) provides identity authentication and permission-management capabilities to deliver baseline identity and access-control support for Agents accessing relevant HUAWEI CLOUD resources. The actual scope of support is subject to the target region, management console and official product documentation.

【**HUAWEI CLOUD Service**】**AgentArts** offers Agent-centric foundational capabilities including platform onboarding, identity verification, access control, quota management and rate limiting, providing security safeguards for Agent invocations of MaaS and associated platform services. Specific functions and applicable coverage are subject to the target region, management console and official product documentation.

5 How HUAWEI CLOUD Assists Clients Use MaaS Securely

5.1 Security Risks and Challenges for Clients Using MaaS

In accordance with the MaaS shared-security responsibility boundary specified in Chapter 3, clients are accountable for the security of content configured, uploaded, developed and consumed by themselves on the MaaS platform. When adopting MaaS, clients shall primarily focus on security risks concerning model selection and application, data processing, ongoing client-side operations, as well as Agent development built upon MaaS. They shall properly configure and leverage the security capabilities offered by HUAWEI CLOUD in line with their business scenarios. This section outlines key security challenges falling within the client's scope of responsibility, while subsequent chapters elaborate on the security features and support available from HUAWEI CLOUD.

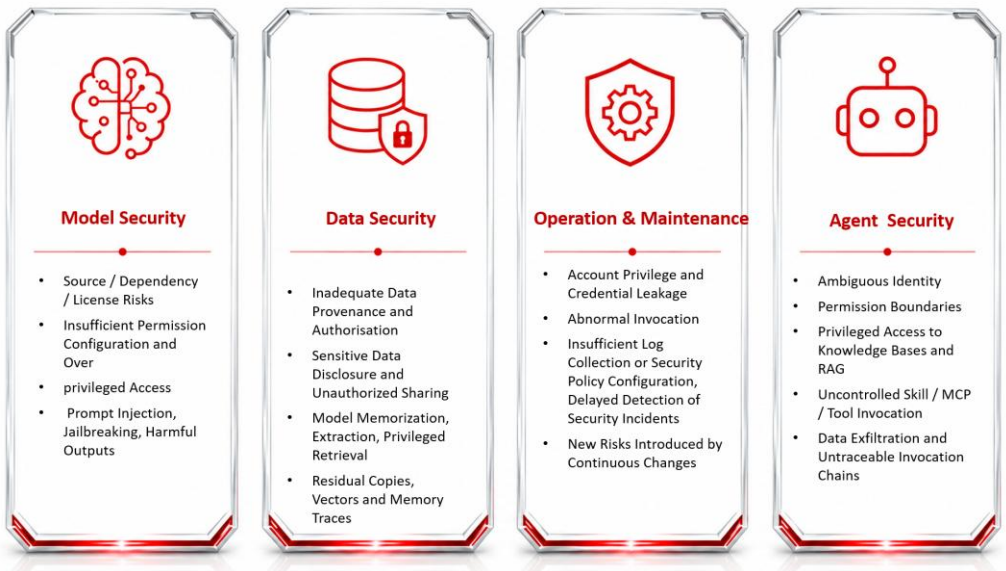

Figure 5-1 Client Security Risk Challenge

Security Risks and Challenges for Models

When clients use HUAWEI CLOUD MaaS, they may select pre-built third-party models or deploy self-defined models according to business requirements. Models vary in capability boundaries, applicable scenarios, version provenance and security risks. Clients shall evaluate model suitability based on business impact and usage patterns, and configure corresponding security policies properly.

During the model selection and onboarding phase, clients may encounter risks including mismatched model capabilities against business requirements, known security vulnerabilities in open-source or custom models, untrusted sources for models and dependent components, and inconsistencies between open-source model license terms and real-world usage scenarios. At the model configuration and application stage, issues such as inadequate security policy settings, over-privileged access, prompt injection, jailbreaking, sensitive-information disclosure, harmful or erroneous outputs may further disrupt client business. As model versions are upgraded and business operations continue, capability boundaries, risk profiles and applicability of models may evolve. Clients are required to continuously assess the security impact of models and their applications.

Phase	Key Risks	Control Focus
Development & Onboarding	Mismatch between model capabilities and business requirements, security flaws in open-source/custom models, provenance or dependency risks, license agreement incompatibility	Model suitability assessment, source and version validation, license review, mandatory security testing
Deployment & Release	Inadequate security policy configuration, over-privileged access, undefined capability boundaries, new risks introduced by version changes	Proper security policy configuration, access permission control, pre-launch validation, additional protective measures where required
Operation & Application	Prompt injection, jailbreaking, sensitive-information disclosure, harmful or erroneous outputs and business misuse	Input-output security controls, content protection, business rule enforcement, manual review and usage boundary management
Change & Decommissioning	Evolving model capabilities, continued use of legacy versions, residual client-owned models or credentials	Post-change re-validation, legacy version deactivation, credential revocation and cleanup of client model assets

Data Security Risks and Challenges

MaaS significantly expands the types, volume and scope of use for client data. Data provided or consumed by clients may enter model workflows as training, fine-tuning and evaluation datasets, or participate in online inference in the form of prompts, conversation contexts and knowledge-base content, and may be further invoked by RAG, Agents and external tools. As data flows continuously across different models, applications and external systems, traditional risks including insufficient data provenance and authorization, sensitive-information disclosure, unauthorized usage, expanded processing scope and incomplete deletion are amplified. Meanwhile, MaaS-specific prominent risks such as data poisoning, model memorization, privilege-escalated retrieval and Agent-driven data exfiltration may emerge.

Clients are primarily exposed to risks associated with data provenance, authorization and intended usage. Evaluation datasets, knowledge bases, user-uploaded files and external data may suffer from unknown sources, inadequate permissions, embedded sensitive information,

poor data quality or malicious poisoning. During parsing, chunking and vectorization of knowledge-base documents, original permission attributes may be lost, rendering originally restricted data accessible to unauthorized parties in subsequent retrieval activities. Accordingly, clients shall place critical focus on legal data sources, authorization boundaries, business purposes, sensitivity classification and input-data security.

Data flows throughout model invocation, knowledge retrieval and application processing. Improper configuration of permissions, storage and transmission on the client side may lead to unauthorized access, accidental sharing or leakage of sensitive data. For RAG and Agent scenarios, misconfigured knowledge-base access rights, persistent memory retention, and invocations via MCP or external tools may further broaden data access coverage and the risk of outbound data transfer. Clients shall implement input-output filtering, data desensitization, access permissions, as well as encryption and flow governance for data within their control, in accordance with data sensitivity levels.

After data usage is completed, residual risks remain for associated copies and derived data. Beyond raw source data, clients shall manage dataset versions, knowledge-base documents, vector indexes, session contexts, memory records and other replicas under their governance. Data that has been used for model training may continue to shape model behavior. Without thorough cleanup aligned with actual data locations, retention periods and business requirements, obsolete data may remain accessible or reusable. Clients shall therefore govern the retention, deletion and retirement of relevant data in compliance with their internal data lifecycle policies.

Operational Security Challenges

Once clients deploy MaaS for ongoing production workloads, they need to continuously monitor the security posture of accounts, model invocations, content outputs and business applications falling within their scope of control. Improper account privilege configuration, leaked access credentials, anomalous model-invocation volumes or sources, and outdated security policies may result in privilege abuse and unintended model calls. Without continuous monitoring for content risks embedded in model outputs and client-built applications, such hazards may accumulate and escalate throughout business operations.

Meanwhile, versions, configurations and business scenarios of client-side models, applications and Agents keep evolving. Pre-launch assessments alone cannot mitigate subsequent emerging risks. Failure to configure essential logging, alerting and security policies tailored to business requirements may delay the detection, investigation and remediation of anomalous invocations, content hazards or application-security incidents. Consequently, clients shall maintain ongoing monitoring over their accounts, service calls and business-security events, promptly respond to privilege abuse, abnormal invocations and content-related risks, and cooperate with HUAWEI CLOUD on necessary investigation, recovery and remediation for incidents involving the MaaS platform.

Agent Security Risks and Challenges

After clients build Agents by connecting enterprise knowledge bases, Memory, Skills, MCP services and business systems on top of HUAWEI CLOUD MaaS, the security boundary expands from model inputs and outputs to identity authorization, knowledge access, tool invocation and business execution. Prompt injection, jailbreaking and biased model capabilities may impair an Agent’s task comprehension and planning, further translating into unintended execution behaviors. Once an Agent is integrated with MCP, Skills and external APIs, malicious tools, over-privileged permissions or dangerous parameters may lead to unauthorized access, erroneous operations or resource abuse.

Meanwhile, sensitive information stored in RAG, Memory and conversation contexts may be subject to privileged retrieval and data exfiltration due to inadequate access controls, malicious content or external calls. Multi-Agent collaboration and third-party components

also increase the complexity of identity, privilege and supply-chain relationships, introducing risks such as permission propagation, uncontrolled invocation chains and untraceable accountability. Accordingly, when building Agents based on MaaS, clients shall priorities the security of Agent identity permissions, knowledge-base access, Skill/MCP and tool invocations, as well as real-world execution actions, and continuously manage associated risks throughout subsequent development, deployment and operation lifecycles.

To address the above-mentioned risks, HUAWEI CLOUD delivers corresponding MaaS security capabilities covering models, data, operations & maintenance, and Agent scenarios to help clients strengthen security governance within their scope of responsibility. Detailed descriptions are provided in Sections 5.2 to 5.5.

5.2 Security for Model Selection and Usage

When clients adopt HUAWEI CLOUD MaaS, they shall select appropriate models in accordance with their business objectives, risk levels and usage patterns, and assume security accountability for configurable model settings, model applications and end-to-end business usage within their control. For third-party or open-source models, clients shall also pay attention to inherent security vulnerabilities, license terms and business applicability; for client-defined custom models, clients are required to govern model provenance, versions and associated components.

Through model specifications, security evaluations, model hosting, access controls, LLM safety guardrails and complementary cloud security capabilities, HUAWEI CLOUD assists clients in identifying model risks, configuring necessary security controls, and consuming MaaS model services in a sustained and secure manner. Platform-native model security guarantees delivered by HUAWEI CLOUD are detailed in Section 4.3. This section focuses on how clients can leverage the above capabilities to fulfil their own model-security responsibilities.

5.2.1 Model Selection and Applicability Management

When selecting models offered by HUAWEI CLOUD MaaS, clients shall comprehensively evaluate the capability boundaries, applicable scenarios and known limitations of models based on business objectives, user coverage, data sensitivity and potential business impacts. Direct deployment for high-risk businesses shall be avoided merely according to general model capabilities or performance metrics. HUAWEI CLOUD may provide information including model descriptions, service statements and security evaluation reports to help clients understand the capabilities and risks of different models; nevertheless, clients shall make the final applicability judgment based on their own business contexts.

For third-party or open-source models provided by MaaS, clients shall also pay attention to model versions, known security vulnerabilities and relevant usage terms. If a model is subject to open-source license requirements, clients shall verify that restrictions on commercial use, geographic coverage, industry sectors and other usage constraints align with real-world business scenarios. For inherent known security risks of models that cannot be eliminated temporarily, clients shall impose necessary usage restrictions or deploy supplementary protective measures according to business impacts.

Clients may perform model selection with reference to model descriptions, security evaluation results, application scopes and known limitations provided by HUAWEI CLOUD. For high-risk businesses such as finance and healthcare, additional business-specific applicability validation shall be conducted. Passing the platform-wide security evaluation shall not be treated as the sole criterion for business go-live. Platform-side security management and

assurance mechanisms implemented by HUAWEI CLOUD for MaaS-hosted models are specified in Section 4.3.

5.2.2 Model Configuration and Security Protection

After finalizing the model to be used, clients shall properly configure model access permissions and security policies in accordance with business risks. HUAWEI CLOUD provides corresponding capabilities including identity authentication, access control, LLM safety guardrails and other cloud-native security features. Clients may implement least-privilege governance tailored to the actual usage patterns of personnel, applications and model services, and reasonably restrict the authorization scope for model invocation, security-policy configuration and other high-privilege operations.

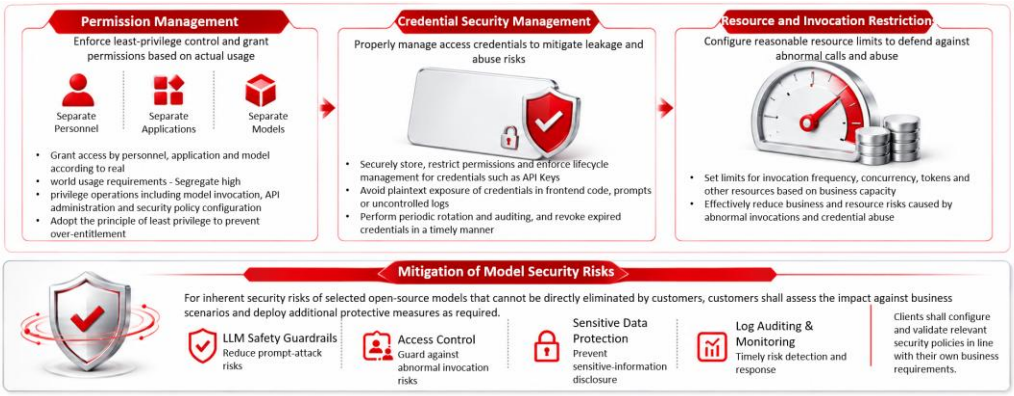

Figure 5-2 Model Configuration and Security Protection

For model-related permissions, clients shall segregate model invocation, API Key administration, security policy configuration and other high-privilege operations to prevent unnecessary entitlements for a single account. Access credentials such as API Keys shall be securely stored, permission-restricted and governed throughout their lifecycle. Credentials must not be placed in plain text within frontend code, prompts or uncontrolled logs. Clients may also configure reasonable limits on invocation frequency, concurrency, tokens and other available resources according to actual business capacity, so as to mitigate business and resource risks arising from abnormal invocations and credential abuse.

With regard to inherent security risks of the selected open-source models that cannot be directly remediated by clients, clients shall assess their impacts against business scenarios and deploy additional protective measures where necessary. HUAWEI CLOUD provides LLM safety guardrails, access control and other applicable security capabilities to help clients reduce risks including prompt injection attacks, anomalous invocations and sensitive-information disclosure. Nonetheless, relevant security policies shall still be configured and validated by clients based on their unique business requirements.

【HUAWEI CLOUD Service】Identity and Access Management (IAM) enables clients to implement authentication and least-privilege management for MaaS-related users, roles and service identities.

【HUAWEI CLOUD Service】MaaS delivers configurable LLM safety guardrails within the inference pipeline. Clients may define corresponding security policies subject to business scenarios and regional service availability. The actual coverage shall be subject to the management console and official product documentation.

5.2.3 Security for Model Application and Ongoing Usage

After models are deployed into real-world business applications, clients shall continuously govern model inputs and outputs, security policies and end-usage patterns in light of application scenarios, service consumers, data sensitivity and potential business impacts. HUAWEI CLOUD MaaS delivers capabilities covering model input-output security, identity and access control to help clients mitigate risks such as prompt injection attacks, harmful content, sensitive-information disclosure and anomalous invocations. Clients are responsible for properly configuring relevant security policies in accordance with business requirements, as well as for the final business consumption of model outputs.

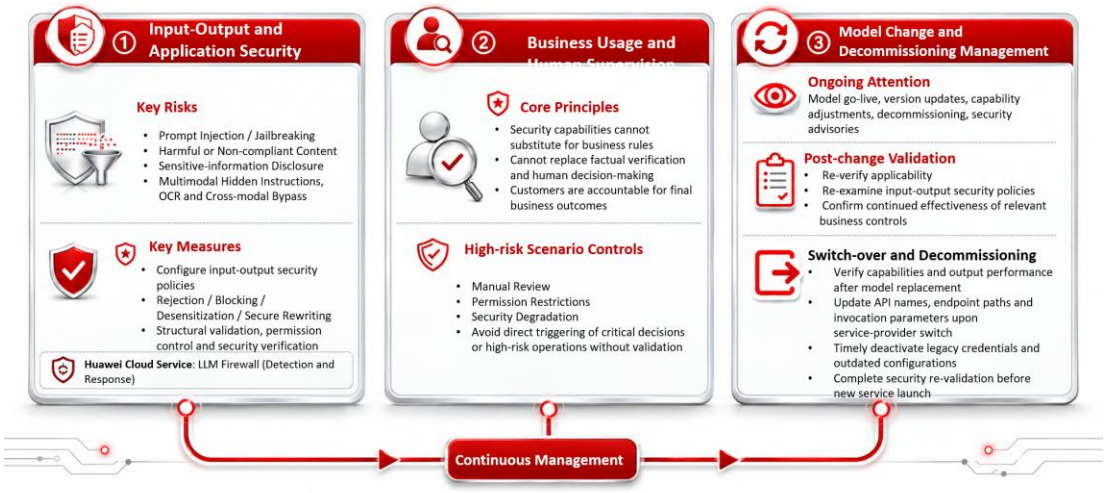

Figure 5-3 Security of Model Application and Usage

5.2.3.1Input-Output and Application Security

Clients shall configure appropriate model input-output security policies based on their application scenarios and risk profiles, with focus on risks including prompt injection, jailbreaking, harmful or non-compliant content, sensitive-information disclosure, as well as hidden instructions, OCR content and cross-modal bypass in multimodal inputs. HUAWEI CLOUD-provided capabilities such as MaaS LLM safety guardrails assist clients in identifying and addressing such risks. Clients shall validate whether policies including rejection, blocking, desensitization and secure rewriting satisfy practical usage requirements against business needs.

For scenarios where model outputs flow further into client business systems, automated workflows or other application components, clients shall implement necessary structural validation, permission control and security verification in light of output types and business impacts. This prevents unvalidated model outputs from being directly executed as high-risk business operations or trusted instructions.

【HUAWEI CLOUD Service】 The **Large Model Firewall** enables detection and mitigation of risks such as prompt attacks, harmful or non-compliant content and sensitive information within model inputs and outputs. It supports multiple minor-language options including Spanish and Portuguese. The actual scope of features is subject to official product documentation.

5.2.3.2Business Usage and Human Supervision

Model security capabilities offered by HUAWEI CLOUD help mitigate security risks throughout model input, inference and output processes, yet they cannot replace clients' own

business rules, factual verification and necessary human decision-making. Clients shall define usage patterns for model outputs in accordance with model capability boundaries and specific business risks, and bear accountability for final business outcomes.

For businesses involving personal safety, property, public interests or other major rights and interests, clients shall deploy necessary controls such as manual review, permission restrictions and security degradation based on actual risks, so as to prevent unvalidated model outputs from directly triggering critical business decisions or high-risk operations. Clients shall also assess and govern the lawful and compliant use of model-generated content in light of their business scenarios and applicable requirements.

5.2.3.3Model Change and Decommissioning Management

Clients shall keep track of model go-live announcements, version updates, capability adjustments, decommissioning notices and relevant security advisories released by HUAWEI CLOUD MaaS, to stay informed about the version and service status of models in use. In case of major model version upgrades, adjustments to capability boundaries or changes in security posture, clients shall re-verify the continued validity of model applicability, input-output security policies and associated business controls based on business impacts, and promptly complete model upgrades or adjust usage plans according to verification results.

When clients need to switch to a different model within HUAWEI CLOUD MaaS, they may update the selected model via application or service configurations, and perform necessary validation on the new model’s capabilities, output performance and existing security policies to ensure continued compliance with business and security requirements after model replacement.

When clients discontinue HUAWEI CLOUD MaaS and migrate to alternative MaaS providers, they shall adjust application invocation configurations for new service interfaces, including updating API names, endpoint paths and relevant invocation parameters. Obsolete HUAWEI CLOUD MaaS access credentials and configurations shall be deactivated in a timely manner to eliminate exposure risks of legacy interfaces or credentials. Prior to launching the new MaaS service, clients shall re-validate model applicability and related security configurations in accordance with internal business and security requirements. Meanwhile, clients shall complete export, backup and migration of relevant data and assets as business needs dictate, and clean up unused resources and access credentials upon migration completion. Requirements for data export, deletion and lifecycle management are specified in Section 5.3.

5.3 Client Data Lifecycle Security

Data security runs through the full lifecycle of MaaS, covering data collection, transmission and storage, processing and manipulation, and data destruction. In accordance with the MaaS shared-security responsibility boundary defined in Chapter 3, clients bear primary responsibility for their own data security. They are accountable for the provenance, authorization and compliant usage of data ingested into MaaS, and shall implement measures such as input-output data filtering, sensitive-data protection, access control, encryption and lifecycle management tailored to business scenarios.

While securing the transmission links and underlying storage of the MaaS platform, HUAWEI CLOUD also delivers security capabilities including identity and permission management, data identification and desensitization, encrypted storage, key management, and audit & traceability. These features help clients fulfil their data-security responsibilities for data usage scenarios across computing infrastructure, models and Agents. Platform-side data

security assurances provided by HUAWEI CLOUD are detailed in Section 4.4. This section focuses on how clients can securely govern and utilise their data.

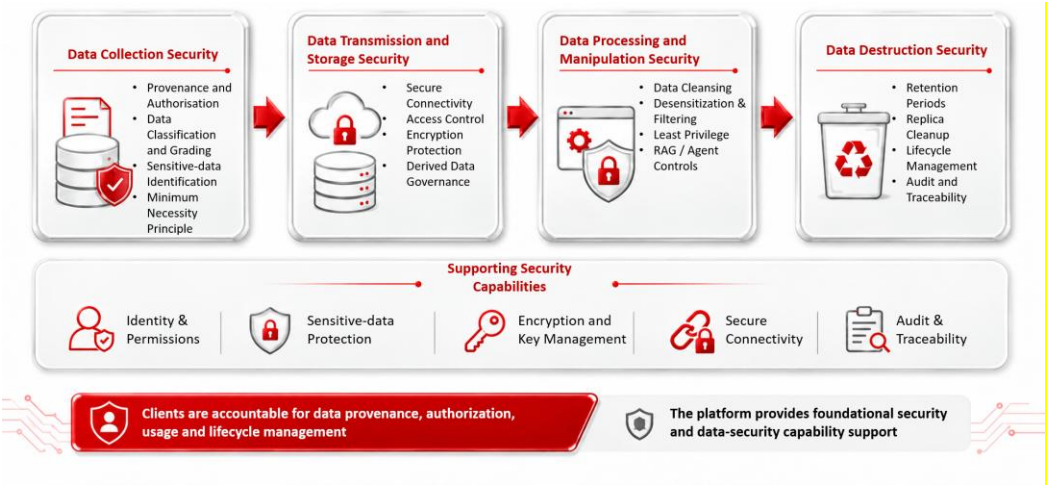

Figure 5-4 Client Data Lifecycle Security

5.3.1 Data Collection Security

Before ingesting or utilizing data in MaaS, clients shall clarify data provenance, authorization scope and intended purposes, and ensure that data acquisition and subsequent usage comply with applicable laws and regulations, contractual agreements and internal data-governance requirements. Data with unknown sources, insufficient authorization, usage beyond defined purposes or embedded unnecessary sensitive information shall be properly processed prior to being fed into MaaS. Where business scenarios permit, personally identifiable information shall not be directly submitted to models. If personal data processing is genuinely required for business purposes, clients shall implement necessary protective measures such as desensitization and de-identification before data is imported into MaaS according to processing purposes, data sensitivity and applicable compliance requirements, and refrain from supplying personal data irrelevant to business objectives to models.

Clients shall perform mandatory data identification, classification & grading and sensitive-information inspection based on data types and sensitivity levels, and limit the scope of data ingested into MaaS following the principle of minimum necessity. For training, evaluation, knowledge-base and other business datasets, clients shall also pay attention to data quality, integrity and trustworthiness, to prevent low-quality, erroneous or malicious data from impairing subsequent model and application operations.

Prior to importing data into MaaS, clients shall restrict the scope of actually uploaded and consumed data under the principle of minimum necessity, and assign corresponding access authorizations for different datasets, knowledge bases and usage scenarios in line with business demands. Data governance and access-control capabilities delivered by HUAWEI CLOUD can facilitate the above-mentioned management work. Nevertheless, clients remain accountable for compliance of data sources, authorization governance, data usage purposes and end-to-end data utilization.

【HUAWEI CLOUD Service】Data Security Center (DSC) supports data asset discovery, data classification & grading, sensitive-data identification and desensitization. It assists clients in detecting sensitive information within datasets intended for model training, evaluation or knowledge bases, mitigating risks of improper sensitive-data usage. Actual feature coverage is subject to regional availability, management console and official product documentation.

【HUAWEI CLOUD Service】**Large Model Firewall** helps clients identify personal data, sensitive information and associated content risks within model inputs and outputs. Clients shall configure and validate corresponding security policies in accordance with business requirements. Actual feature coverage is subject to regional availability, management console and official product documentation.

【HUAWEI CLOUD Service】**Identity and Access Management (IAM)** delivers identity authentication and permission-management capabilities, enabling clients to govern access scopes for MaaS-related data, knowledge bases and resources. Actual feature coverage is subject to official product documentation.

5.3.2 Data Transmission and Storage Security

During MaaS consumption, client data may continuously flow among client business systems, HUAWEI CLOUD MaaS services, storage services and associated applications, and be transmitted and persisted in various forms including business data, evaluation datasets, knowledge-base data, model contexts, vector data, session records and logs. HUAWEI CLOUD assumes responsibility for baseline transmission and storage security for MaaS data transmission links and hosted storage environments under its control, as detailed in Section 4.4. For network connections and other storage resources created, configured and controlled by clients themselves, clients shall properly configure access permissions, encryption policies, keys, data replicas and retention periods according to data sensitivity levels.

In terms of data transmission, clients shall select appropriate secure connection modes based on their deployment architectures and constrain the scope of data in transit. For data transmitted across networks or to external systems, avoid passing sensitive information irrelevant to business purposes. Within client-controlled storage environments, access permissions and encryption measures shall be configured according to data classification levels, with proper governance over data versions, replicas and lifecycles.

For cache, vector data, session contexts, Memory, logs and other derived data generated during model inference, clients shall identify their actual storage locations and access scopes, and define corresponding retention periods and protection requirements aligned with business needs. This prevents scenarios where original data is well-protected while derived data remains exposed for extended periods.

【HUAWEI CLOUD Service】**Object Storage Service (OBS)** can be used to store client-managed datasets, evaluation sets and other MaaS-related data. For OBS buckets created and controlled by clients, clients are responsible for configuring bucket policies, access permissions, encryption, version control and lifecycle rules. Platform-side storage security responsibilities for hosted storage managed by the HUAWEI CLOUD MaaS platform are specified in Section 4.4. Actual feature coverage shall be subject to official product documentation.

【HUAWEI CLOUD Service】**Data Encryption Workshop (DEW)** delivers key management and encryption-related capabilities. For data and storage resources under client control, clients may configure corresponding encryption measures based on data sensitivity and business requirements, and govern access and lifecycles of client-side keys. Actual feature coverage shall be subject to official product documentation.

【HUAWEI CLOUD Service】At the network layer, secure connectivity for diverse business scenarios is provided via the HTTPS protocol, or services such as **VPC** endpoints, **VPN** and **Direct Connect**. Clients may select appropriate data transmission methods according to their real-world deployment architectures. Actual feature coverage shall be subject to regional availability and official product documentation.

5.3.3 Data Processing and Manipulation Security

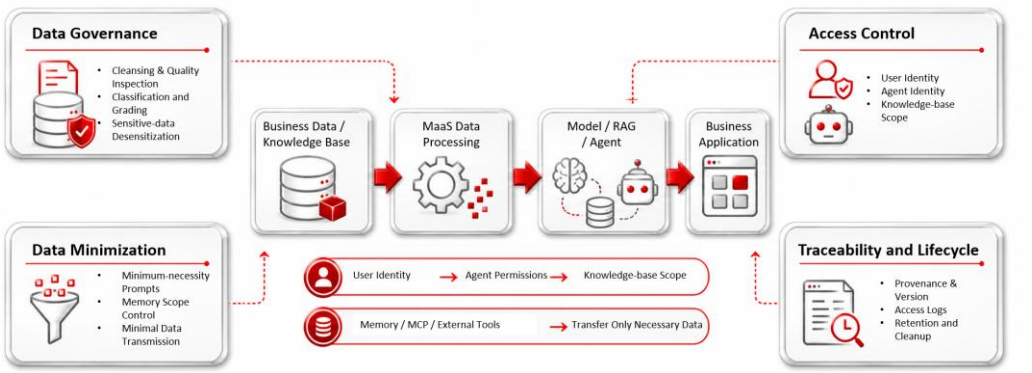

Figure 5-5 Data Processing and Manipulation Security

After data is ingested into MaaS, clients shall perform requisite data cleansing and governance in compliance with data-quality and regulatory requirements. This includes identifying and rectifying erroneous, duplicate, anomalous or inappropriate data, amending missing or incomplete information, and applying filtering, desensitization or other protective measures for sensitive information, personal data and non-compliant content, so as to mitigate risks of improper data propagating to subsequent model processing and application workflows.

For scenarios requiring data annotation, clients shall govern annotation scope, label content and annotation outcomes, and retain corresponding version and process records. Prior to data being used for model evaluation, inference or other MaaS applications, clients shall conduct necessary data security and quality assessments tailored to real-world scenarios, with focus on data integrity and accuracy, data bias, content compliance and annotation correctness. Data failing to meet requirements shall be cleansed, corrected or discontinued from usage.

Once data enters the operational phase, clients shall constrain data access and usage scope following the principle of minimum necessity, and deploy permission restrictions, desensitization and mandatory input-output data security filters according to data sensitivity. For RAG and Agent scenarios, knowledge-base retrieval scope shall be restricted based on the identities of end-users and Agents, to prevent unauthorized cross-access to data from different departments, users or classification levels. For Memory, MCP and external tool invocations, the scope of transmitted data shall be limited to avoid full-context payloads or unnecessary sensitive information being passed to third parties.

Clients shall also maintain adequate records of data provenance, versions, processing activities and access events to enable traceability of data usage. Mechanisms such as access control and desensitization shall be leveraged to reduce direct exposure of raw sensitive data. Requirements for data retention periods, expiry-driven cleanup and relevant audit controls are further specified in Section 5.3.4.

【HUAWEI CLOUD Service】Data Security Center (DSC) assists clients in identifying sensitive information within data processed by MaaS, and implements measures including data classification & grading, sensitive-data identification and desensitization for real-world scenarios. Actual feature coverage is subject to official product documentation.

【HUAWEI CLOUD Service】Large Model Firewall helps clients detect sensitive information and associated content risks within model inputs and outputs. Clients shall configure and validate relevant security policies in accordance with business requirements. Actual feature coverage is subject to regional availability, management console and official product documentation.

【HUAWEI CLOUD Service】**AgentArts** delivers management capabilities for knowledge bases, Agents and associated resources, helping clients govern access to knowledge data and the scope of data consumption within Agent applications. Actual feature coverage is subject to official product documentation.

【HUAWEI CLOUD Service】**DataArts Studio** provides capabilities such as data quality management to support clients with data cleansing, quality inspection and related data governance activities. Actual feature coverage is subject to regional availability, management console and official product documentation.

【HUAWEI CLOUD Service】**ModelArts** offers dataset management, annotation and relevant data-processing capabilities, delivering platform support for client-side data preparation, annotation and quality management. Actual feature coverage is subject to regional availability, management console and official product documentation.

5.3.4 Data Destruction Security

During MaaS usage, raw data, dataset versions, knowledge bases, vector indexes, session contexts, Memory and other derived data may be generated. Before changing MaaS usage patterns, migrating services or unsubscribing from services, clients shall identify data and associated assets to be retained based on business continuity, data retention and compliance requirements. They shall complete necessary data export, backup or migration leveraging available product capabilities, and verify the integrity and usability of exported data.

For client-controlled data, models, knowledge bases and related resources, HUAWEI CLOUD provides corresponding capabilities for data access, download, export or migration via MaaS and associated cloud services to facilitate client data preparation prior to service decommissioning. Clients shall complete required data retention and migration before unsubscription, and timely deactivate obsolete API Keys, access permissions and relevant configurations.

After finishing data migration or business exit, clients shall clean up data and replicas within their scope of control in accordance with data usage purposes, compliance requirements and sensitivity levels. HUAWEI CLOUD cleans or deletes client data and platform-side data that are no longer required under its control, in compliance with applicable service rules and data lifecycle mechanisms. For OBS buckets, databases, knowledge bases, logs and other associated resources created or controlled by clients, cleanup operations shall be performed by clients according to actual data locations. The specific exportable data types, formats, retention periods and deletion mechanisms are subject to the product capabilities, service terms and official product documentation applicable to clients.

【HUAWEI CLOUD Service】**Object Storage Service (OBS)** supports object lifecycle and version management capabilities. Clients may configure lifecycle rules based on their own data retention requirements to execute expiry-based processing for storage data under client control. Actual feature coverage is subject to official product documentation.

【HUAWEI CLOUD Service】**ModelArts** and **AgentArts** deliver relevant dataset, knowledge-base and resource management capabilities, assisting clients in implementing corresponding lifecycle management according to actual data locations. Actual feature coverage is subject to official product documentation.

【HUAWEI CLOUD Service】**Cloud Trace Service (CTS)** records key management operations of relevant cloud resources, providing support for audit and traceability of data lifecycle management. The scope of auditable events is subject to official product documentation.

5.4 Operational and Maintenance Security

After clients adopt MaaS for ongoing production-grade business operations, they shall continuously operate and maintain accounts, model invocations, security configurations, content risks and the security status of business applications within their scope of control. Clients are responsible for continuous monitoring of account anomalies, model calls, content risks and business security events on the client side. They shall configure logs, alerts and security policies in line with business requirements, and promptly remediate permission abuse, abnormal invocations, content risks and application-related security incidents.

While ensuring the sustained secure operation of the MaaS platform, HUAWEI CLOUD provides capabilities including asset and invocation information, identity and permission management, audit logs, monitoring & alerting, and content security. These help clients perform AI asset management, risk monitoring, security incident response and continuous compliance management within their areas of responsibility. Platform-side operational and maintenance security assurances from HUAWEI CLOUD are detailed in Section 4.5. This section focuses on how clients can leverage relevant capabilities to use MaaS securely on an ongoing basis.

On this foundation, clients may utilize HUAWEI CLOUD-provided AI security posture management and intelligent security operation capabilities to correlate and manage MaaS-related assets, risks, logs, alerts and security events. Security agents can be adopted to improve efficiency in asset identification, risk detection, threat analysis and response-strategy generation, establishing a continuous operational mechanism featuring “visible assets, traceable risks, trackable incidents and closed-loop remediation”.

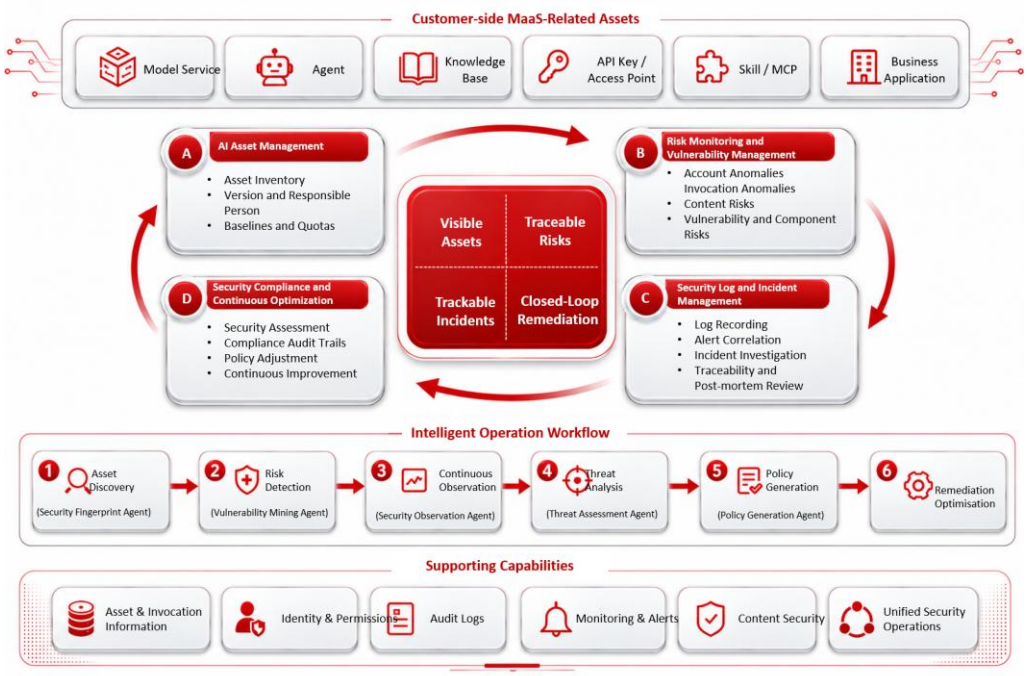

Figure 5-6 Client-Side Operational and Maintenance Security

5.4.1 AI Asset Management

After adopting MaaS, clients shall establish a continuous management mechanism for AI-related assets selected, configured, created and consumed by themselves, and clarify asset responsible persons, business purposes, versions, permissions, operating statuses and correlation relationships. As clients keep adding model services, access points, datasets,

knowledge bases and business applications, relevant assets and configurations will change dynamically. Clients shall keep abreast of asset creation, modification and decommissioning, so as to avoid unmanaged assets that are unassigned, long-term idle or utilized beyond expected scopes.

Clients shall build their own MaaS asset inventory, covering at minimum model services and their versions, service access points, API Keys and other access credentials. Knowledge bases, Agents, Skill/MCP and relevant external connections shall be further included based on real-world usage scenarios. Clients may record asset names, business purposes, versions or statuses, responsible entities, usage scopes and correlation relationships to underpin subsequent risk inspection and change management.

To improve the efficiency of MaaS asset discovery and correlation governance, clients may leverage HUAWEI CLOUD-provided security fingerprint agent capabilities to assist in identifying AI assets including model services, Agents, knowledge bases, vector databases, Skill/MCP, APIs and relevant connections. Asset fingerprints are generated based on asset types, versions, operating statuses, permission configurations and invocation characteristics. Invocation and dependency relationships among assets further help detect unmanaged, idle or anomalous assets, which shall be timely incorporated into the client's MaaS asset inventory and subsequent security management scope. The security fingerprint agent referenced in the solution is designed for automatic AI asset discovery, attribute collection and asset fingerprint inventory establishment.

Clients shall regularly reconcile their asset inventory against actual resources, configurations and invocation records on the MaaS platform. Asset information shall be synchronously updated upon model version changes, addition or deactivation of service access points, creation or deletion of API Keys, and adjustments to knowledge-base or business-application configurations. Obsolete model services, access points, API Keys and other client-controlled assets shall be deactivated or cleaned up in a timely manner to eliminate exposure risks posed by idle assets and legacy credentials.

For critical MaaS assets, clients shall set normal-usage baselines and configure reasonable invocation quotas according to business capacity. For high-privilege API Keys, knowledge bases involving sensitive data, high-impact business applications and other high-risk assets, clients may, in light of business.

【HUAWEI CLOUD Service】Cloud Trace Service (CTS) records key management operations such as IAM authorization changes, model service activation or deactivation, access-point modifications, and API Key creation or deletion within its supported scope. It supports clients in tracing MaaS asset changes and reconciling asset inventories. The scope of auditable events is subject to regional availability, management console and official product documentation.

5.4.2 Risk Monitoring and Vulnerability Management

After clients deploy MaaS for ongoing production-grade operations, they shall continuously monitor the security status of accounts, model invocations, content outputs and business applications within their scope of control, rather than relying solely on pre-launch checks to assess business security. Clients shall define necessary monitoring metrics and risk baselines based on normal business usage patterns, and promptly identify and investigate behaviors deviating from normal conditions.

In terms of accounts and credentials, clients shall watch for risks including anomalous logins, permission modifications, abnormal API Key usage and other indicators suggestive of account or credential abuse. For model invocations, abnormal patterns can be identified via dimensions such as invocation volume, frequency, source, time window, model and access

point. Attention shall be paid to security and cost risks triggered by sudden traffic surges, unexpected sources or anomalous model calls.

Regarding content and model utilization, clients shall keep monitoring risks such as prompt injection, jailbreaking, harmful or non-compliant content, and sensitive-information leakage in alignment with business scenarios. Monitoring shall be implemented with HUAWEI CLOUD-provided model input-output security capabilities and client-side application security policies. For self-built Agents, third-party component integration and model-enabled business applications, clients shall also watch for Agent behavioral drift, abnormal tool invocations, unauthorized knowledge-base access, third-party component anomalies and related business security incidents.

Building upon continuous risk monitoring, clients may leverage HUAWEI CLOUD intelligent security operation capabilities to improve risk detection and analysis efficiency. The vulnerability-mining agent assists in risk detection and security inspection for client-controlled AI applications, Agents, plugins and associated components. The security-observation agent continuously correlates model invocations, Agent behaviors, tool calls, data access and other security signals to detect abnormal activities deviating from normal-usage baselines. The threat-assessment agent aggregates, deduplicates and cross-correlates multi-source alerts to reconstruct risk chains and determine impact scope and risk severity. The policy-generation agent produces recommendations for security policies, permission adjustments, risk remediation or emergency response based on assessment outcomes, improving client risk-handling efficiency. Clients retain the responsibility for final judgement and implementation of risks and countermeasures in light of their business scenarios and actual control scope. The four capabilities above correspond to the intelligent operation workflow of “Risk Detection — Continuous Observation — Threat Assessment — Policy Generation” in the reference solution.

For application code, Agents, Skill/MCP, third-party plugins and relevant components under client control, clients shall keep track of vulnerabilities and security updates, and perform timely upgrades, patching, access restriction or other necessary remediation according to vulnerability severity, exposure status and business impact. Vulnerability detection and remediation for the HUAWEI CLOUD MaaS platform and hosted components are managed by HUAWEI CLOUD in accordance with the responsibilities defined in Chapter 4.

Upon detection of account anomalies, invocation irregularities, content risks or business security incidents, clients shall conduct preliminary analysis using asset, identity, invocation and business information. Timely countermeasures including access restriction, permission tightening, API Key rotation, invocation throttling and security-policy adjustment shall be adopted to prevent risk escalation.

【HUAWEI CLOUD Service】SecMaster: It delivers cloud-native unified security operations. Based on the full-life-cycle closed-loop risk-management framework, it provides four agents for vulnerability mining, security observation, threat assessment and policy generation. It helps clients proactively identify risks and continuously monitor model invocations and business access. Actual feature coverage is subject to regional availability, management console and official product documentation.

5.4.3 Security Log and Incident Management

Clients shall perform adequate logging and management for MaaS-related operations, model invocations, security detections and business application logs within their scope of control, in accordance with business risks, audit requirements and incident-response needs. Reasonable log retention periods, access permissions and protective measures shall be configured. Logs may contain account information, prompts, model invocation payloads or other sensitive data, for which clients shall adopt corresponding protective measures based on practical conditions.

During daily operations, clients shall pay attention to logs and security alerts related to account logins, permission and API-Key modifications, model-service invocations, knowledge-base access, and the runtime status of Agents and business applications. In case of anomalous logins, sudden invocation-volume surges, unexpected access sources, content risks or business security incidents, clients shall conduct investigations using relevant logs, alerts and asset information to identify incident origins, impact scopes and response requirements.

For security incidents requiring in-depth analysis, clients may leverage unified security-operation platforms and intelligent-analysis capabilities to correlate logs, alerts, asset records and risk information. This facilitates reconstruction of incident chains and impact-scope evaluation, followed by remediation and post-mortem reviews as appropriate. For cross-boundary incidents involving the HUAWEI CLOUD platform, client applications or third-party components, both parties shall conduct necessary information sharing and coordinated response in compliance with the responsibility boundaries and actual control scopes defined in Chapter 3.

Clients shall retain sufficient logs, alerts, investigation records and remediation records to support subsequent incident post-mortems, security assessments and compliance audits. Access to logs shall follow the principle of minimum privilege with proper safeguards, preventing sensitive logs from becoming a security risk in themselves.

【HUAWEI CLOUD Service】Cloud Trace Service (CTS) records key MaaS management-plane operations including IAM authorization changes, model service activation or deactivation, access-point modifications, and API Key creation or deletion within its supported scope. It assists clients with operation auditing and incident traceability. The scope of auditable events is subject to regional availability, management console and official product documentation.

【HUAWEI CLOUD Service】Object Storage Service (OBS) can be used for centralized storage of client-dumped audit logs. Clients may configure access permissions, retention periods and corresponding protective measures in accordance with their security and compliance requirements. Actual feature coverage is subject to official product documentation.

【HUAWEI CLOUD Service】Cloud Eye Service (CES) enables clients to configure monitoring and alerts for relevant service metrics, and notify designated responsible persons of critical anomalies in a timely manner. Supported metrics, alerting methods and regional coverage are subject to official product documentation.

【HUAWEI CLOUD Service】SecMaster allows clients to integrate MaaS-related security signals into a unified security-operation view according to actual access scope. It supports multi-source alert correlation, risk assessment, incident investigation and coordinated response. Actual feature coverage is subject to regional availability, management console and official product documentation.

5.4.4 Security Compliance and Continuous Optimization

Clients shall continuously identify cybersecurity, data security, personal-information protection, generative-AI and industry regulatory requirements applicable to their businesses, considering their country or region of operation, industry sector, data types, user groups and specific MaaS application scenarios. Relevant requirements shall be implemented within security measures under client control, covering model utilization, data processing, account permissions, log management, content security and business applications.

Clients shall conduct regular security assessments for MaaS usage based on their internal risk and compliance requirements. Mandatory reassessments shall be performed against practical risks upon major model changes, significant adjustments to business scenarios, modifications

to critical permissions or data-processing approaches, and occurrence of major security incidents. Assessments shall focus on model applicability, compliant data usage, access permissions, security policies, log-alert configurations, incident response, and the effectiveness of control measures for client applications and Agents. Identified issues shall be remediated and re-verified.

In accordance with business and audit requirements, clients shall retain necessary security-assessment records, configuration-change logs, incident records, critical-security configurations and supporting evidence, including applicable security-assessment documentation, key-security-configuration records, audit logs, alert and incident-response records, as well as relevant product certifications and service materials provided by HUAWEI CLOUD.

Clients may further leverage alerts, risk intelligence and remediation experience accumulated throughout risk monitoring, incident investigation and post-mortem reviews to iteratively adjust monitoring baselines, security policies and operational procedures. This enables security controls to evolve alongside changing MaaS business conditions and threat landscapes, forming a continuous-improvement closed loop of “Monitoring — Assessment — Response — Optimization”.

【HUAWEI CLOUD Service】Cloud Trace Service (CTS) records key management operations of cloud resources within its supported scope for clients, assisting them in generating evidence for operation audits and security assessments. The scope of auditable events is subject to official product documentation.

5.5 Agent Security

After clients build Agents based on MaaS, model capabilities are further connected to knowledge bases, Memory, Skill, MCP, APIs and business systems. AI applications are extended from content generation to tool invocation and business execution. In accordance with the MaaS shared-security responsibility boundary defined in Chapter 3, clients are responsible for end-to-end lifecycle security covering secure development, deployment and operation of Agent applications. Clients shall focus on managing Agent identity permissions, knowledge access, connections to tools and external systems, business execution behaviors and runtime auditing, to mitigate risks such as prompt injection, excessive privileges, unauthorized retrieval, malicious tool invocation and unintended business operations.

HUAWEI CLOUD provides baseline authentication, access control, quotas, rate limiting, together with relevant data, model and security operation capabilities to help clients implement Agent security controls. Baseline platform security capabilities are detailed in Section 4.6. This section mainly describes security requirements to be implemented by clients when building and using Agents on HUAWEI CLOUD MaaS.

5.5.1 Agent Identity and Access Control

When building Agents, clients shall differentiate among identities including natural persons, business applications, service accounts and Agents. Independent, identifiable and traceable identities shall be configured for Agents based on real-world business requirements. Long-term sharing of high-privilege accounts or access credentials across multiple Agents or business scenarios shall be avoided. When Agents access models, knowledge bases, APIs, MCP and other tools, required permissions shall be configured following the principle of least privilege, so that Agents can only access resources and operations necessary to complete current business tasks.

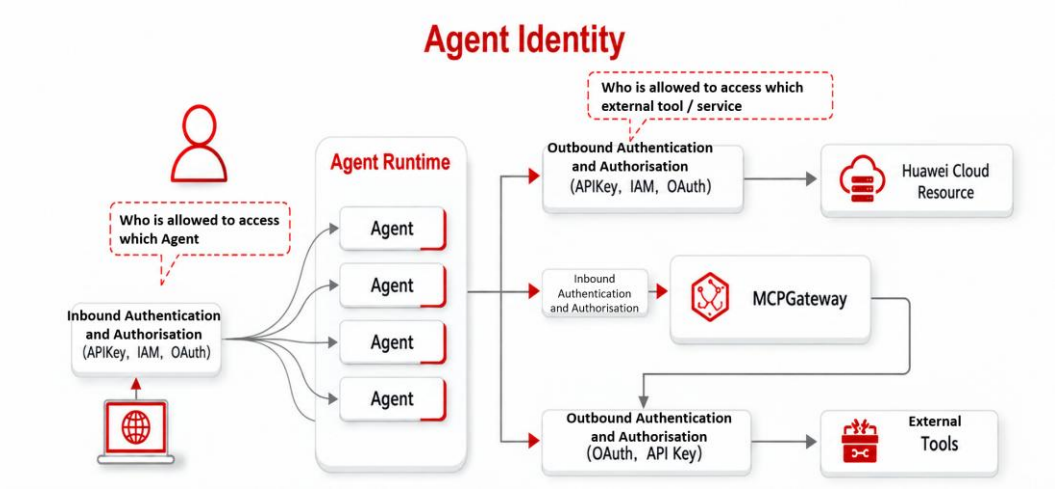

Figure 5-7 Agent Identity and Access Control

Clients shall establish permission boundaries based on the Agent's business role, data sensitivity and operational impact, and grant necessary authorizations separately for model invocation, knowledge-base retrieval, tool usage and external-system access. For high-privilege Agents and Agents capable of high-risk operations such as data modification, transactions, approvals, message sending and system configuration, their authorization scope shall be further restricted. User confirmation, manual approval or other privilege-escalation mechanisms shall be configured according to business risks.

Agent identities, permissions, access credentials and associated resources shall be subject to lifecycle management. When there are changes to an Agent's purpose, responsible person, connected systems or permissions, permissions shall be reassessed and adjusted in a timely manner. When an Agent is decommissioned or no longer uses relevant tools and external systems, corresponding permissions and credentials shall be revoked promptly to mitigate risks arising from persistent legacy authorizations.

【HUAWEI CLOUD Service】Identity and Access Management (IAM) assists clients in implementing identity authentication and permission control for relevant HUAWEI CLOUD resources, and configures the scope of resources accessible to Agents following the principle of least privilege. Actual feature coverage is subject to official product documentation.

【HUAWEI CLOUD Service】AgentArts provides management capabilities for Agent identities and permissions of associated resources, helping clients govern access relationships between Agents and platform resources. Actual feature coverage is subject to regional availability, management console and official product documentation.

5.5.2 RAG and Enterprise Knowledge Base Security

When clients ingest enterprise documents, business data or external content into Agent knowledge bases, they shall clarify knowledge sources, applicable scopes and access permissions, and retain original permission attributes such as user, department, project, tenant and classification level wherever possible. During RAG retrieval, access control shall be enforced based on both the identity of the current end-user and the Agent identity, ensuring that the Agent only retrieves knowledge content accessible to the requesting subject, and mitigating risks of unauthorized cross-user, cross-department or cross-classification retrieval.

For web pages, user-uploaded files and other external content, clients shall also pay attention to impacts of prompt injection, malicious instructions and untrusted data on Agent behaviors,

preventing external content from being misinterpreted as system instructions that further trigger knowledge leakage or unintended operations. When knowledge-base content, permissions or authorization status change, relevant retrieval permissions and indexes shall be synchronously updated in a timely manner to avoid persistence of invalid authorizations.

When Agents transmit data to models and external services using knowledge-base content, RAG results, session context or Memory, clients shall limit data scope following the principle of least necessity, and avoid passing full context or unnecessary sensitive information to irrelevant models, tools or third-party services. More detailed requirements for data sources, classification, grading, desensitization and lifecycle management are specified in Section 5.3.

【HUAWEI CLOUD Service】AgentArts provides capabilities for knowledge-base ingestion, document processing, knowledge retrieval and associated resource management, assisting clients in building Agent knowledge-access workflows. Specific permission-control functions and supported scope are subject to regional availability, management console and official product documentation.

【HUAWEI CLOUD Service】Data Security Center (DSC) assists clients in sensitive-data identification, classification-and-grading and desensitization for knowledge bases and relevant data. Actual feature coverage is subject to official product documentation.

5.5.3 Skill、MCP Plugin Security

When connecting Skills, MCP services, plugins, APIs or other external tools to Agents, clients shall select trusted components based on business purposes, and understand the sources, functional scopes, access permissions and external dependencies of such components prior to usage. For self-developed, self-deployed or self-integrated Skills, MCP services and third-party tools, clients are responsible for their secure utilization and the security management of relevant components and dependencies.

Clients shall restrict data, interfaces and business systems accessible to Agents via tools in accordance with the principle of least privilege, and avoid providing high-privilege credentials beyond actual task requirements to Agents. For tools capable of modifying business data, invoking high-privilege interfaces, executing code, initiating transactions or sending outbound messages, the scope of invokable methods, parameters and resources shall be further constrained, and necessary user confirmation or manual approval shall be added according to business risks.

Clients shall also watch for malicious instructions or untrusted content that may be contained in tool descriptions, MCP return payloads and responses from third-party services, and prevent Agents from executing external content directly as highly-trusted instructions. For Skills, MCP services and external tools with changed sources, version upgrades, expanded permissions or identified security vulnerabilities, clients shall reassess their risks and timely adjust, restrict or discontinue their usage。

【HUAWEI CLOUD Service】AgentArts provides connectivity and management capabilities for Agents with relevant tools, knowledge and platform resources, helping clients govern resource-invocation relationships within Agent applications. Supported Skills, MCP, gateway and permission-control capabilities are subject to regional availability, management console and official product documentation.

5.5.4 Runtime Isolation and Auditing

A key distinction between Agents and ordinary model invocations is that model outputs may be further utilized for task planning, tool invocation and real-world business operations. Clients shall treat model outputs as non-deterministic results requiring further verification,

and shall not directly adopt them as high-risk business operations or trusted execution instructions without adequate validation and permission controls.

Before an Agent performs business operations, clients shall implement necessary parameter validation, permission checks and security verification according to operation types and potential impacts. For high-risk operations involving personal safety, property, modification of critical data, outbound transmission, transactions, approvals, system configuration or other significant interests, mechanisms such as user confirmation, manual approval, privilege escalation or termination shall be added based on actual risks. This prevents Agents from completing major or irreversible operations relying solely on a single model output.

Clients shall also constrain the scope of executable code, tools, network access and resource consumption for Agents. Restricted runtime modes shall be adopted for dynamic code, untrusted tasks or high-privilege tools. Invocation counts, concurrency, timeouts and other resource limits shall be configured in line with business capacity to avoid abnormal resource consumption or business disruptions caused by anomalous tasks, recursive calls or tool abuse.

During runtime, clients shall log necessary information including Agent identity, model invocations, knowledge retrieval, tool calls, key parameters, user confirmation and execution results according to business risks, to support anomalous-behavior detection, invocation-chain analysis and business-incident traceability. Access permissions and retention periods shall be configured for relevant logs based on data sensitivity, and unnecessary sensitive information shall not be logged.

Upon detection of Agent permission abuse, abnormal tool invocations, unauthorized knowledge-base access, unintended business execution or other application security incidents, clients shall promptly take countermeasures such as invocation throttling, credential revocation, permission tightening, suspension of relevant tools or Agent services, and conduct investigation and remediation following their own incident-response procedures. For incidents involving the HUAWEI CLOUD MaaS platform, clients shall cooperate with HUAWEI CLOUD for necessary investigation and recovery in accordance with the shared-responsibility boundary defined in Chapter 3.

Agents together with their knowledge bases, tool connections, permissions and access credentials shall be managed under application lifecycle management. When material changes occur to an Agent’s business purpose, model, knowledge base, Skill/MCP or external systems, clients shall re-examine relevant permissions and security controls. When an Agent is decommissioned, identities, access credentials and associated permissions shall be revoked in a timely manner. Knowledge bases, Memory and other data assets under client control shall be processed in compliance with the data-lifecycle requirements set forth in Section 5.3.

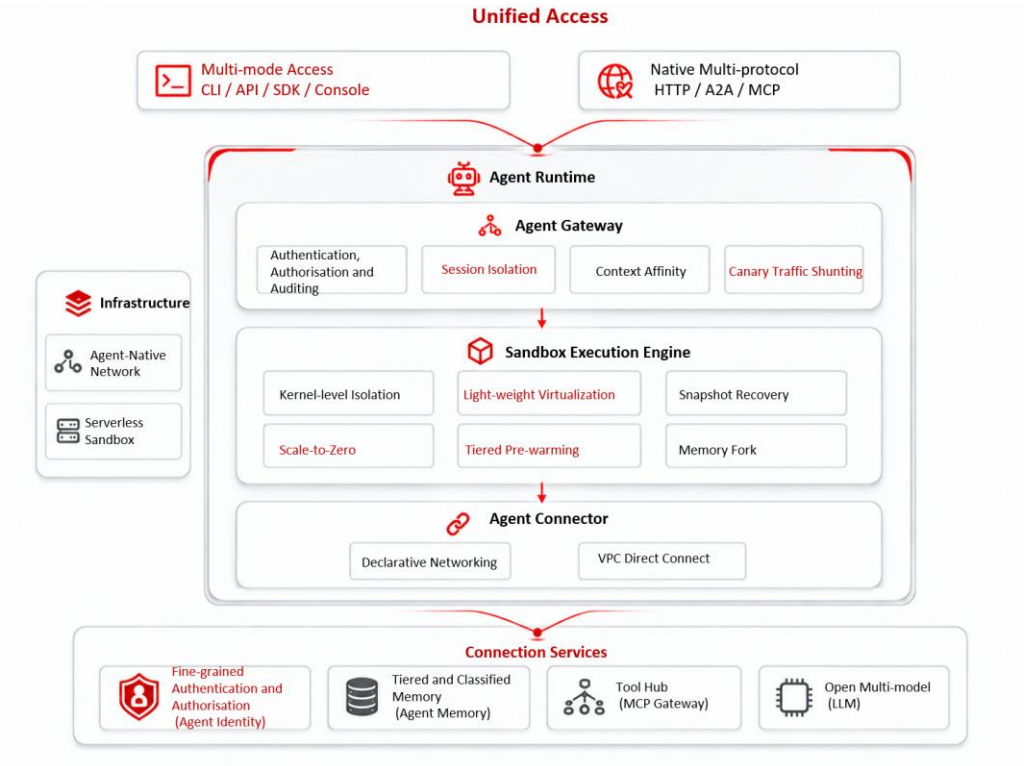

Figure 5-8 Agent Runtime Isolation

【**HUAWEI CLOUD Service**】**AgentArts** delivers Agent runtime environment and relevant management capabilities. Clients may configure identity verification, access control, resource limits and other runtime-security capabilities within the applicable scope. Actual feature coverage is subject to regional availability, management console and official product documentation.

【**HUAWEI CLOUD Service**】**Cloud Trace Service (CTS)**, together with associated logging and monitoring capabilities, supports clients in resource-operation traceability and security operations within its supported scope. The scope of available logs, events and monitoring is subject to official product documentation.

5.5.5 Intent and Execution Security

When clients build and deploy Agents based on MaaS, in addition to identity authentication, permission control and resource isolation, they shall treat Agents as digital employees with autonomous operational capabilities. Leveraging HUAWEI CLOUD Agent Guard, clients shall establish behavioral control baselines aligned with job responsibilities, and enforce dynamic security protection for key runtime behaviors and execution intents, so as to prevent role overreach and high-risk operations caused by autonomous planning, semantic misunderstanding or task-execution deviation.

Clients shall define digital job roles and behavioral baselines for Agents in accordance with enterprise organizational structures and job responsibilities, and specify accessible data, invocable tools and permissible business operations. For example, finance-oriented Agents shall not access core code repositories, and coding-oriented Agents shall not perform financial transactions. This prevents logical privilege escalation arising when Agents break job-responsibility boundaries during autonomous planning.

Clients shall also define clear data and instruction boundaries for Agents, and explicitly distinguish system instructions, user instructions and external data. External documents, web

content, tool return outputs and other third-party data shall be treated as untrusted or low-trust data. Such content shall not obtain the semantic status of system-level instructions via content injection, nor shall it override existing task objectives, permission constraints or security policies. During multi-round task execution, Agents shall continuously verify the consistency between current sub-goals and user-authorized tasks, avoiding deviation from original tasks induced by external data or anomalous context.

Clients may use HUAWEI CLOUD Agent Guard to detect runtime behaviors and execution intents of Agents. By correlating information such as task context, user identity, knowledge retrieval, tool invocations and target resources, high-risk behaviors including data erasure, privilege escalation and exfiltration of sensitive information can be identified, with real-time interception or restriction enforced before actual execution. For high-risk operations such as data deletion, financial transactions, permission modification, approval workflows and outbound message delivery, manual confirmation, manual approval or dynamic circuit-breaker mechanisms shall be configured in alignment with enterprise business-approval processes.

When an Agent deviates from its job-behavior baseline or triggers high-risk intents, relevant tasks shall be suspended, tool invocations restricted, or operations handed over to human-controlled checkpoints in a timely manner, forming a dynamic security closed-loop where machines provide defense and humans retain ultimate control.

6 MaaS Security Practices

The preceding chapters have introduced HUAWEI CLOUD MaaS security capabilities from dimensions including computing infrastructure, models, data, and operations-maintenance. This chapter further elaborates how these security capabilities can be combined and applied from the perspective of real-world client scenarios. First, it describes how security boundaries can be extended to knowledge bases, tools and business systems when clients build Agents using MaaS model capabilities. Second, it explains how financial institutions implement security controls addressing data, model, platform and business risks when using MaaS under a heavily-regulated environment.

6.1 Industry Case: MaaS Security Practice for a Financial Institution

6.1.1 Client Requirements

As large-language models are gradually adopted in core business scenarios such as intelligent claim settlement, employee office assistants and code development, financial institutions have raised higher requirements for data transmission security, network boundary isolation and data usage scope when consuming MaaS services. Although traditional internet-based access to MaaS services can leverage technologies such as HTTPS to secure transmission, for security-sensitive clients including financial and state-owned enterprises, routing business data over public-network links may still fail to satisfy their internal security management and risk-control requirements.

Against this backdrop, the client intended to utilize cloud-hosted large-model capabilities while further mitigating the risk of business-data exposure on the public internet. The goal was to realize intranet transmission of client data between the enterprise data center and MaaS services. Combined with security measures including network isolation, access control and transmission encryption, a more secure and controllable large-model access environment was to be established.

6.1.2 MaaS Security Practice

To address the client's security requirements, HUAWEI CLOUD built a MaaS intranet access solution based on VPC and Cloud Connect (CC). The client can deploy business systems for intelligent claim settlement, office assistants, code development and other workloads within its on-premises IDC or cloud-hosted VPC. Business networks are securely interconnected with the MaaS service network via the access chain "Client IDC — Cloud Connect (CC) — Client VPC — MaaS VPC". Model invocations and associated business data are transmitted over in-cloud networks, reducing reliance on public-network access links.

In terms of network security, VPC delivers logical isolation between the client's business environment and other network environments. Security groups and access-control capabilities are applied to restrict the access scope of MaaS services. During data transmission, dedicated network links and encryption mechanisms further safeguard data in transit and mitigate risks

of eavesdropping, tampering and unauthorized access. As a result, a multi-layer protection framework of “Network Isolation + Intranet Access + Link Encryption” is implemented.

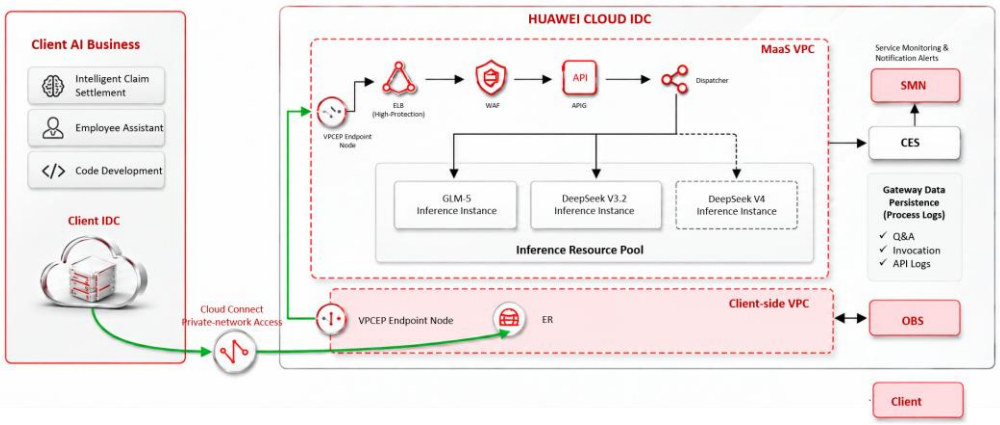

Figure 6-1 MaaS Example of Intranet Transmission Solution

6.1.3 Client Value

With the above-mentioned solution, clients can leverage large-model capabilities while extending security controls to the access links of model services. On one hand, dedicated intranet access to MaaS enables business data to be transmitted without traversing the public internet and reduces public-network exposure. On the other hand, end-to-end transmission encryption, network isolation, access control, content review and prompt-security protection are combined to enforce security protection for the full lifecycle of model invocations.

Meanwhile, addressing financial clients’ high priority on data protection, measures including prohibition of data usage for model training, service-side data-processing controls and relevant security-compliance certifications further clarify data-usage boundaries. A multi-layer security protection system covering “access, transmission, data, content and model” is established in general, helping financial institutions apply MaaS capabilities to business scenarios such as intelligent claim settlement, internal office work and R&D in a more secure and controllable manner.

6.2 MaaS Security Configuration Best Practices

This chapter applies to clients who build and run AI applications using HUAWEI CLOUD MaaS services, covering network security, identity management, data protection, asset management, as well as logging and threat detection. All configurations shall be implemented in accordance with actual business conditions, region-available capabilities and organizational compliance requirements.

6.2.1 Network security

6.2.1.1Invocation Channel Selection

1-1 | Prioritise intranet link establishment via VPCEP

When business systems are deployed within HUAWEI CLOUD VPC, access MaaS through VPC Endpoint (VPCEP) as the first choice, so that invocation traffic is transmitted over the intranet to reduce public-network exposure. For scenarios outside the VPC or across clouds, extend intranet reachability using Direct Connect (DC) or VPN.

Reference: **https://support.huaweicloud.com/bestpractice-vpcep/vpcep_05_1001.html**

1-2 | Use public-network HTTPS as an alternative channel and minimize exposure

If intranet access is unavailable, public-network HTTPS endpoints may be used. Refer to 1-3 and 1-4 for requirements on transmission protocols and server-certificate verification; refer to 1-5 for requirements on backend proxy and frontend credential protection. The channel priority order is: VPCEP intranet, intranet extended by Direct Connect or VPN, public-network HTTPS (after egress convergence is completed). Front-ends shall not directly invoke model endpoints over public-network HTTPS.

6.2.1.2Transmission Encryption

1-3 | Mandatory use of TLS 1.2 or later

All invocations shall use TLS 1.2 or higher. Enable TLS 1.3 preferentially if conditions permit. Plain-text HTTP shall be disabled, and protocol or cipher-suite downgrade shall be avoided.

1-4 | Strict server-side certificate validation

Client SDKs or HTTP libraries shall enable certificate and hostname verification. verify=False shall not be configured, nor shall TLS errors be ignored. When outbound traffic passes through proxies or gateways, ensure there is no unauthorized or uncontrolled TLS termination and re-encryption in intermediate links.

6.2.1.3Invocation Link Convergence

1-5 | Invoke via backend proxy; front-ends shall not connect directly to model endpoints

Front-ends shall not hold credentials and send requests to MaaS endpoints directly. Requests shall be proxied by back-ends, BFF or gateways; front-ends only communicate with their own back-end services. For specific requirements on secure credential hosting and periodic rotation, see 2-5 and 2-6 respectively.

1-6 | Centralize outbound egress points

All outbound invocations shall be converged to a small number of fixed egress points, such as unified NAT, outbound proxies or dedicated subnets, to facilitate centralized security control and monitoring. Fixed egress IP addresses can be used as the basis for configuring source-IP whitelists for API Keys; see 1-8 for detailed whitelist requirements.

1-7 | Outbound access control (minimize destinations and ports)

Only destination addresses and ports required for accessing MaaS endpoints shall be permitted on the egress side, e.g., HTTPS port 443 or VPCEP intranet addresses. Outbound connections irrelevant to business operations shall be restricted.

6.2.1.4Access-Source Control

1-8 | Configure source-IP whitelisting for API Keys

When creating an API Key, bind permitted source IP addresses for invocations to the whitelist. Refer to 1-6 for requirements on centralized outbound egress points. It is recommended to populate the whitelist with the fixed public IP addresses of unified backend outbound egress points, so that each API Key maps to its corresponding egress. Even if an API Key is compromised, sources outside the whitelist cannot initiate invocations using the credential. This control is therefore recommended as a baseline requirement.

Reference: **https://support.huaweicloud.com/intl/zh-cn/model-call-maas/model-call-049.html**

1-9 | Supplementary source-IP and geographic controls at self-built entry points

Source-IP and geographic restrictions for end-users shall be enforced on self-built WAF or API Gateway. This control applies to the “end-user-to-client-application” access path. The source-IP whitelisting specified in 1-8 applies to the “client-backend-to-MaaS” invocation path.

6.2.1.5Protection for Self-built AI Application Entry Points

1-10 | WAF-based entry-point protection

Deploy WAF at the entry points of self-built applications according to business exposure surface, to defend against injection attacks, malicious scraping and malicious request payloads. Configure reasonable limits for large-size payloads and over-long contexts.

Reference: **https://support.huaweicloud.com/waf/index.html**

1-11 | Anti-DDoS/DDoS High-Protection

Enable baseline Anti-DDoS protection for public-network entry points. For critical services, configure enhanced DDoS protection based on risk assessment and business-continuity requirements.

Reference: **https://support.huaweicloud.com/productdesc-aad/aad_01_0001.html**

1-12 | APIG Unified Entry

Self-built applications may expose services externally through API Gateway (APIG), to centrally implement TLS enforcement, rate limiting, concurrency control, access control and observability management.

Reference: **https://support.huaweicloud.com/apig/index.html**

No.	Check Item	Control Responsibility	Compliance Criteria
1-1	VPCEP Intranet Access	To be enabled and configured by client	Business workloads within VPC invoke via VPCEP intranet; traffic does not traverse public network
1-2	Public-network HTTPS Alternative	Provided by platform by default	HTTPS is used when intranet access is unavailable; endpoints and credentials are not exposed on front-end
1-3	TLS 1.2 or Higher Version	Provided by platform by default	Client enables TLS 1.2 or higher; protocol downgrade is prohibited
1-4	Certificate Validation	Implemented by client	Certificate and hostname verification are enabled; verify=False is not configured
1-5	No Direct Front-end Connection	Implemented by client	Front-end only communicates with client-owned backend; credentials are not stored on front-end
1-6	Egress Convergence	Implemented by client	Outbound invocations are converged to a small number of fixed egress points
1-7	Outbound Access Control	Implemented by client	Only destinations and ports required for MaaS endpoints are permitted at egress
1-8	Source-IP Whitelist for API Key	To be enabled and configured by client	Fixed egress IPs are bound upon API Key creation and included in security baseline
1-9	Source / Geolocation Control for Self-owned Entry	Implemented by client	End-user source restrictions are enforced via WAF / APIG
1-10	WAF for Self-owned Entry	To be enabled and configured by client	Self-owned AI entry points are integrated with WAF
1-11	DDoS Protection	Provided by platform by default + to be enabled and configured by client	Baseline protection is enabled for public-network entry points; enhanced protection is configured for critical services based on risk
1-12	APIG Unified Entry	To be enabled and configured by client	External services are exposed through APIG with rate-limiting and unified governance capabilities

Table 6-1 Cybersecurity Checklist

6.2.2 Identity Management

6.2.2.1Least Privilege and Separation of Duties

2-1 | Isolate API Key issuance and deletion operations with fine-grained permissions

Enforce separation of duties using MaaS fine-grained API-Key permissions: Grant modelarts:apikey:create (create or edit) and modelarts:apikey:delete (delete) only to a small group of API Key administrators. Ordinary users shall not be assigned such permissions; they may only consume already-issued API Keys and shall have no authority to issue or revoke keys on their own.。

2-2 | Isolate credential boundaries by application / environment

Use mutually independent API Keys for different applications and environments (development, test, production), each configured with separate permission scopes and source constraints. Refer to 2-3 and 2-4 for specific requirements on API-Key permission scope and source-IP whitelists. Such containment measures limit the impact scope of credential compromise to a single application and a single environment.

6.2.2.2API Key Full-Lifecycle Management

2-3 | Set custom access scope upon issuance

When creating an API Key, select “custom access” rather than “full access” as the preferred permission scope. Restrict the API Key to only access model services or endpoints required for the corresponding business, mitigating risks of unauthorized access to other model services in case of credential leakage.

2-4 | Bind source-IP whitelist upon issuance

Bind a source-IP whitelist at the time of API-Key creation, and configure fixed IP addresses of unified backend outbound egress points as permitted sources. Refer to 1-8 for configuration requirements for source-IP whitelists.

2-5 | Place API Keys under secure hosting immediately after creation

The plaintext API Key is displayed only at creation. Immediately import it into a credential-management service such as Cloud Secret Management Service (CSMS) or Data Encryption Workshop (DEW), and verify that applications can securely retrieve it. Plaintext credentials shall not be stored or transmitted via screenshots, instant-messaging tools, ordinary text files or other uncontrolled channels.

2-6 | Periodic rotation with canary-based replacement to avoid service disruption

Define fixed rotation cycles according to risk levels. Perform immediate rotation upon suspected compromise, personnel changes or permission adjustments. Complete replacement following the sequence: create a new key, perform canary traffic shift, confirm zero traffic on the old key, then delete the old key, so as to reduce business-disruption risks. If the platform does not support automatic expiry, clients shall proactively implement rotation by deleting and recreating keys.

2-7 | Leakage emergency-response plan

Document formal response procedures in advance. Upon suspected API-Key compromise (e.g., receiving abnormal-invocation alerts generated per detection rules specified in Section 5-3), sequentially revoke the compromised API Key, issue a new API Key with appropriately tightened permission scope and source-IP constraints, switch business traffic, investigate leakage paths and conduct post-incident reviews. Conduct regular drills for this procedure.

No.	Check Item	Control Responsibility	Compliance Criteria
2-1	Isolation of API Key Issuance and Deletion Permissions	To be enabled and configured by client	Create and delete permissions are granted only to designated API Key administrators
2-2	Application and Environment Isolation	Implemented by client	Separate API Keys are used for each application and each environment
2-3	Custom Access Scope	To be enabled and configured by client	Production-environment API Keys can only access model services or endpoints required by business workloads
2-4	Source-IP Whitelist	To be enabled and configured by client	Bind fixed egress IP addresses upon key issuance; see 1-8 for configuration requirements
2-5	Secure Storage of Plaintext Credentials	Provided by platform by default + Implemented by client	Plaintext is visible only at creation; import to secure hosting immediately; no plaintext credential exposure
2-6	Periodic Rotation	Implemented by client	Defined rotation cycle; canary-based replacement without service disruption
2-7	Compromise Emergency Response	Implemented by client	Documented formal procedure with regular drills

Table 6-2 Identity Management Checklist

6.2.3 Data Protection

6.2.3.1Data Classification and Minimization

3-1 | Classify and grade data within input prompts

Establish data-classification-and-grading and access-admission rules on the application side. Public data may be invoked according to business requirements. Internal and sensitive data, such as personal information, identity-document or bank-card data, keys and passwords, undisclosed source code and core commercial data, shall in principle not be directly fed into models. If processing is genuinely required for business purposes, implement desensitization or de-identification in accordance with the data-protection requirements specified in 3-2, and complete necessary authorization and compliance assessment.

3-2 | Desensitization / de-identification of sensitive fields prior to invocation

Desensitize, mask, pseudonymize or replace identifiable personal-information fields and highly-sensitive fields with placeholders before requests leave the client backend. For instance, real names or identity numbers may be replaced by tokens, which are restored on the client side after responses are received. Desensitization shall be completed within the client backend and shall not rely on platform-side processing.

6.2.3.2Protection of Data in Transit

3-3 | End-to-end encrypted transmission; plain-text and weak protocols prohibited

MaaS API endpoints utilize TLS encryption. Refer to 1-1 to 1-4 for detailed requirements on transmission-channel selection, TLS versions and server-certificate validation. Data shall remain encrypted throughout transmission. Plain-text protocols and weak protocols such as TLS 1.0 and TLS 1.1 shall be disabled.

No.	Check Item	Control Responsibility	Compliance Criteria
3-1	Established and application-enforced data-grading rules for prompt-eligible data	Implemented by client	Rule documentation exists; sensitive-level data is blocked by default
3-2	Sensitive fields are desensitized on client backend before submission; no reliance on platform-side processing	Implemented by client	No plain-text PII or credentials found in sampled requests
3-3	All invocations use HTTPS with certificate validation enabled; weak protocols are not permitted	To be enabled and configured by client	Packet capture / configuration verification confirms TLS 1.2+ and certificate validation is turned on

Table 6-3 Data Protection Checklist

6.2.4 Asset Management

6.2.4.1Asset Inventory and Ledger

4-1 | Establish MaaS asset ledger based on platform portals

Perform asset inventory by leveraging console portals including model-service or subscription lists, endpoint configurations, API Key lists, invocation monitoring and statistics, and build a unified client-side ledger. Recommended ledger fields include asset type, unique identifier, owning team or application, activation or creation time, status and person-in-charge. Non-sensitive attributes such as API Key permission scope, source-IP whitelist, tags and creation time may be recorded, whereas plain-text credentials shall not be logged.

4-2 | Periodically reconcile ledger against actual platform assets

Compare platform asset lists with the client-side ledger on a regular basis (e.g., monthly) to identify model services or API Keys missing from the ledger. Such assets may be leftover resources after personnel changes or uncleaned temporary test resources. Upon detection, bring them under ledger management, or decommission model services and delete API Keys in accordance with asset-disposal procedures.

6.2.4.2Quota, Usage and Cost Visibility

4-3 | Establish usage baselines via platform invocation monitoring and quota views

Use platform-provided invocation-monitoring, statistics and quota views to establish normal-usage baselines by dimensions such as model service, endpoint or API Key, covering invocation volume, frequency and quota consumption. Usage baselines shall be updated periodically and serve as thresholds for anomaly detection.

4-4 | Configure reasonable upper quota limits for assets to prevent uncontrolled consumption

Configure appropriate upper quota limits for model services or invocations based on business capacity and peak-demand requirements, mitigating risks of runaway invocations and abnormal cost growth caused by application malfunctions or API Key compromise.

No.	Check Item	Control Responsibility	Compliance Criteria
4-1	Asset ledger covering model services / endpoints / API Keys has been established	To be enabled and configured by client + Implemented by client	Ledger exists with complete fields and corresponds to platform portals
4-2	Periodically reconcile records against the platform to identify and dispose of unmanaged assets	To be enabled and configured by client + Implemented by client	Reconciliation records exist; no assets exist outside ledger management
4-3	Usage baselines for each asset are obtained (via platform monitoring / quota views)	To be enabled and configured by client	Dimension-based usage-baseline records are available
4-4	Reasonable upper quota limits are configured for critical assets	To be enabled and configured by client	Quotas are set and aligned with business volume

Table 6-4 Asset Management Checklist

6.2.5 Logging and Threat Detection

6.2.5.1Log Collection and Retention

5-1 | Enable CTS cloud audit covering key MaaS management-plane events

Enable HUAWEI CLOUD Cloud Trace Service (CTS). Based on the actual scope of auditable events, cover key management-plane events including IAM authorization changes, model-service activation or de-subscription, endpoint modifications, API Key creation or deletion, and delegate operations, to provide audit evidence for identity and asset management.

Reference: **https://support.huaweicloud.com/cts/index.html**

5-2 | Independent retention of dumped logs with configured retention periods and tamper-resistance protection

Dump CTS events to a standalone Object Storage Service (OBS) bucket for centralized persistence. Configure retention periods and tamper-resistance controls that comply with organizational compliance requirements. Enforce separation of duties to prevent operators from deleting or modifying audit logs related to their own operations.

6.2.5.2Threat Detection

5-3 | Detect credential-abuse signals against usage baselines

Take the usage baselines specified in 4-3 as a reference, and configure detection rules for the following deviations (using platform invocation monitoring or Cloud Eye Service (CES) metrics; rules and thresholds are defined by the client). First, abnormal invocation volume or frequency, such as sudden traffic surges, quota exhaustion or abnormally high call rates. Second, abnormal sources, e.g., invocation attempts originating outside the IP whitelist

specified in 1-8. Third, abnormal time windows or patterns, such as regular invocations during off-business hours, or anomalous combinations of models and endpoints.

5-4 | Enable content-security review to detect abuse at input / output layers

For risks including malicious content generation, prompt injection and induced leakage of sensitive information, enable and ingest risk signals from MaaS-native content-security review capabilities if available in the target region. For scenarios not covered by the platform, clients shall implement supplementary validation for inputs and outputs on the application side.

5-5 | Ingest platform-side protection logs (WAF / Anti-DDoS) to supplement threat-monitoring views

WAF and Anti-DDoS protections deployed per 1-10 and 1-11 generate attack-blocking logs. Ingest such logs into the overall threat view and perform correlation analysis against invocation-layer anomalies. For example, identify sources triggering both WAF blocks and abnormal invocations to improve threat-investigation accuracy.

6.2.5.3Alerting, Response and Orchestration

5-6 | Configure cloud-monitoring alerts for real-time notification of responsible personnel upon critical threats

Use HUAWEI CLOUD Cloud Eye Service (CES) to configure alerts for thresholds, sudden mutations and invocations originating outside whitelists against key metrics and events specified in 5-3. Notify responsible persons recorded in the asset ledger in real-time via SMS, email or Webhook. Alert information shall clearly map to specific assets and their corresponding owners.

Reference: **https://support.huaweicloud.com/ces/index.html**

5-7 | Develop threat-response playbooks and link with remedial actions

Formulate and conduct drills for threat-response playbooks, and link alerts with existing remediation measures. When anomalies are detected, revoke or tighten permissions for suspected-compromised API Keys in accordance with the credential-leakage emergency-response procedure defined in 2-7, adjust quotas temporarily following the quota-management requirements in 4-4, and harden whitelists and entry-point controls per the access-source-control requirements in 1-8 and 1-9. Afterwards, investigate leakage paths, conduct root-cause post-mortems and update relevant security policies.

5-8 | Periodically review logs and alerts for continuous optimization of control strategies

Periodically audit-review audit logs and alert records, assess the effectiveness of detection rules, and optimize configuration strategies specified in Clause 1 through 4 accordingly to build a closed-loop for continuous improvement.

No.	Check Item	Control Responsibility	Compliance Criteria
5-1	CTS is enabled and covers key MaaS management-plane events	To be enabled and configured by client	CTS event list includes IAM / service / key modification events
5-2	Logs are independently dumped with retention period, tamper-resistance, and operators cannot delete their own logs	To be enabled and configured by client + Implemented by client	Dedicated dump bucket in place; retention period and permission isolation

No.	Check Item	Control Responsibility	Compliance Criteria
			are properly configured
5-3	Credential-abuse detection rules are established for abnormal invocation volume, sources and time windows	To be enabled and configured by client + Implemented by client	Detection rules are defined and based on usage baselines established per 4-3
5-4	Platform-side content review enabled (if available) plus application-side content validation	Provided by platform by default + To be enabled and configured by client + Implemented by client	Content-risk signals are collected and remediated
5-5	Platform-side WAF / Anti-DDoS logs are ingested into threat-monitoring view	Provided by platform by default	Protection logs are used for correlation analysis
5-6	Alerts are configured for critical threats to notify designated responsible persons in real time	To be enabled and configured by client	CES alert rules exist; notification channels are validated to work
5-7	Threat-response playbook available, integrated with revocation / quota / whitelist controls	Implemented by client	Documented playbook exists and has been exercised via drills
5-8	Periodic review of logs and alerts for continuous optimization of security configurations	Implemented by client	Review records and continuous-improvement records for security configurations are available

Table 6-5 Logging and Threat Detection Checklist

7 Authoritative Certifications

7.1 Authoritative Security Certifications for Large-Model Offerings

Leveraging core strengths including full-featured capabilities, high performance, high reliability and superior user experience, HUAWEI CLOUD MaaS has obtained certifications and assessments issued by multiple authoritative institutions. These credentials verify its industry-leading performance across security, trustworthiness, performance and service capability dimensions, and serve as references for stakeholder solution selection.

Certification Name	Certification Description
Inspection Certificate for Trustworthiness Requirements of Large-Scale Pre-trained Models	A "Trusted AI" -series assessment conducted by the China Academy of Information and Communications Technology (CAICT) in accordance with *Evaluation Methods for Technologies and Applications of Large-Scale Pre-trained Models*. Benefiting from abundant industry applications and outstanding security-and-trustworthiness capabilities, HUAWEI CLOUD large-model product is the first large-model product to obtain the Excellent-Level (Level 5) rating. It also achieved an Excellent-Grade (4+ level) score in compliance verification against artificial-intelligence technical and product capability standards.
Maturity-Level Certificate for Large-Model Service Capabilities	*Artificial Intelligence — Large Models — Part 3: Assessment of Service Capability Maturity* is the first national-level standard for large-model To-B service-capability domain, jointly compiled by the China Electronics Standardization Institute (CESI) together with more than 50 enterprises and institutions. HUAWEI CLOUD is among the first batch of enterprises passing all three test items of the large-model service-capability maturity assessment organized by CESI, attaining the Collaborative-Optimization maturity level.
Trusted AI Assessment Certificate (High-Quality Token Service	The "Trusted AI-Token Service Quality Assessment" carried out by CAICT based on *Capability Requirements for Artificial-Intelligence Token Service Quality*. Assessed HUAWEI CLOUD MaaS (Model-as-a-Service) meets token-service-quality requirements for enterprise-grade general-purpose scenarios and

Certification Name	Certification Description
Series - Service Quality)	was awarded one of the first-batch "Trusted AI-High-Quality Token Service" assessment certificates.
Self-assessment Certificate for Compliance with National Standards on Artificial-Intelligence Security	The *Cybersecurity Technology — Basic Security Requirements for Generative-AI Services* was released by the National Cybersecurity Standardization Technical Committee. HUAWEI CLOUD participated in the pilot self-assessment for generative-AI-service security. Assessment results verify that its model-security capabilities satisfy standard requirements, and the corresponding certificate is issued by CESI.

7.2 Authoritative Certifications Obtained by HUAWEI CLOUD

HUAWEI CLOUD Computing Technologies Co., Ltd. has obtained a number of domestic and international management-system certifications, aiming to deliver a robust compliance assurance system covering key domains including AI governance, data security and privacy compliance, which demonstrates HUAWEI CLOUD's AI compliance capabilities.

Certification Name	Certification Description
ISO/IEC 42001	To address the rise of artificial intelligence and its associated challenges, the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) have released the world′ s first AI-management-system standard, ISO/IEC 42001. HUAWEI CLOUD has obtained ISO/IEC 42001 certification, and the certificate bears the internationally-recognized ANAB accreditation mark. This certification covers HUAWEI CLOUD AI products, spanning R&D, testing, operation-and-maintenance and service-support for businesses including AI models, AI cloud platforms and AI application assistants. Attaining ISO/IEC 42001 certification serves as key evidence that an enterprise′ s AI governance, compliance and sustainable-development practices have reached internationally-advanced levels. It fully demonstrates that Huawei′ s AI governance is fully aligned with international standards and sets a benchmark for responsible AI governance.
Data Security Management Certification	Data Security Management Certification (DSM Certification) is conducted by the State Administration for Market Regulation and the Cyberspace Administration of China based on the national standard *Information Security Technology — Security Requirements for Network-Data Processing* (GB/T 41479) and other relevant standards and specifications. This certification demonstrates that HUAWEI CLOUD complies with stringent standard requirements for data-processing activities including data collection, storage, usage, processing, transmission, provision and disclosure.

Certification Name	Certification Description
Data Management Capability Maturity Level Certification	DCMM (Data Management Capability Maturity Model) corresponds to the national standard *Data Management Capability Maturity Assessment Model* (GB/T 36073-2018), the first officially-released national standard in China's data-management domain. HUAWEI CLOUD has achieved Level 5, the highest maturity level, which proves its leading position in data platforms and data-management services.

In addition, HUAWEI CLOUD has also obtained authoritative certifications and audits including ISO/IEC 27001, ISO/IEC 27701, ISO/IEC 27018, ISO/IEC 29151, and ISO/IEC 22301. More certifications obtained by HUAWEI CLOUD can be viewed via the **Compliance Center**- **Compliance Certification** on HUAWEI CLOUD's official website. Apart from these third-party certifications, the Compliance Center also provides **Country/Region-Specific Compliance Guidance** and **Industry-Specific Compliance Guidance** to help resolve issues you may encounter in compliance-related work.

HUAWEI CLOUD is committed to building secure and trustworthy services. Through assessments and certifications issued by independent authoritative third-party institutions, we continuously deliver dependable intelligent products and services with optimal experience for clients, and jointly promote the beneficial development of AI.

8 Appendix

8.1 HUAWEI CLOUD MaaS Security Capability List

Capability Domain	Capability	Built-in MaaS Capability (Default Security Capability)	Client-Optional Capability (Advanced Security Capability)	Security Product / Service
Asset and Access Security	AI Asset Identification and Unified Management	√	√	AgentArts
	Console and API Identity Authentication	√		
	Fine-Grained Permission Control	√	√	IAM
	Delegated Authorization and Temporary Credentials	√	√	IAM ModelArts Delegated Authorization
Data Protection	Static Encryption for Sensitive Personal Information		√	DEW
	Data-in-Transit Encryption	√		
	Sensitive Data Desensitization	√	√	DSC
	Model and Asset Integrity Verification	√		
	Tenant Data Isolation	√		
Platform and	Web Application Attack Protection	√		

Capability Domain	Capability	Built-in MaaS Capability (Default Security Capability)	Client-Optional Capability (Advanced Security Capability)	Security Product / Service
Runtime Protection	Host and Computing Environment Protection	√		
	Vulnerability Scanning and Remediation	√		
	High-Volume-Attack Protection	√		
	Database Security Protection	√		
	Model Input-Output Content Security	√	√	Large Model Firewall
	Resource Isolation and Interface Access Restriction	√		
Security Operations	Audit and Traceability for Key Operations	√	√	CTS
	Online-Service and Model Monitoring & Alerting	√	√	CES
	Threat Awareness and Incident Response	√	√	SecMaster
	Fault Detection and Automatic Recovery	√		
	Service Security Updates	√		

8.2 FAQ

- **Question 1: What main security risks should I pay attention to when using HUAWEI CLOUD MaaS?**

  Answer: When using MaaS, you shall focus on security risks within your scope of control, including model selection and application, data processing, accounts and access credentials, model input and output, business applications, and Agent operation. Examples include mismatch between model capabilities and business scenarios, prompt injection, jailbreaking,

harmful or erroneous outputs, sensitive-information leakage, API Key compromise or abnormal invocations, as well as unauthorized retrieval, tool misuse and unintended business execution in RAG and Agent scenarios. For security risks within HUAWEI CLOUD's control scope such as chips, servers, underlying cloud resources, the MaaS platform and Huawei-Cloud-hosted components, HUAWEI CLOUD provides corresponding security guarantees in accordance with the MaaS shared security responsibility boundary.

- **Question 2: What are the respective security responsibilities of HUAWEI CLOUD and myself when using MaaS?**

  Answer: MaaS security responsibilities are divided mainly based on the actual control, visibility and operability of both parties over relevant resources, data and businesses. HUAWEI CLOUD is responsible for the security of cloud and MaaS hosted services themselves, including the computing infrastructure, MaaS platform and runtime environment for hosted models, platform-side data transmission and underlying storage, platform security operations, and basic Agent security components. You are responsible for your own configuration and usage activities on MaaS, including data source, authorization and compliant usage, model selection and business-applicability assessment, account permission and API Key management, security-policy configuration, model input-output, and security of self-built applications and Agents.

- **Question 3: What security guarantees does the HUAWEI CLOUD MaaS platform itself provide?**

  Answer: HUAWEI CLOUD has built security assurance mechanisms covering computing power, models, data and continuous operations throughout the MaaS service chain, including trusted computing and resource isolation, tenant isolation, secure transmission and storage protection, identity and access control, model security evaluation and runtime protection, model input-output security, log auditing, risk monitoring, vulnerability management, and security incident response. Meanwhile, you may configure or leverage relevant security capabilities such as IAM, LLM Firewall, DSC, DEW and CTS according to business requirements. Actual available capabilities are subject to the target region, models in use, console and official product documentation.

- **Question 4: How does HUAWEI CLOUD help me mitigate security risks during model usage?**

  Answer: HUAWEI CLOUD implements corresponding model security management for MaaS model services provided or hosted by itself, including model source and supply-chain inspection, security evaluation, integrity protection, deployment and runtime protection, security defect remediation, and version management. On the client-usage side, you may select models suitable for your business scenarios with reference to HUAWEI CLOUD-provided model descriptions, security evaluation results and known limitations. Capabilities including identity and access control, LLM security guardrails and invocation limits can be adopted to mitigate risks such as prompt injection, jailbreaking, harmful content, sensitive-information leakage and abnormal invocations. For high-risk businesses such as finance and healthcare, you shall conduct necessary applicability verification, factual checking and human supervision based on your own business rules; general platform-side security evaluations shall not replace business-oriented security judgements. For more AI security solutions, please refer to the ***AI Security Solution Whitepaper***.

- **Question 5: What precautions should I take when using third-party models or third-party model APIs in the Model Gallery?**

  Answer: For third-party models in the Model Gallery, you shall understand and comply with the corresponding license agreements, usage restrictions and other third-party requirements, and assess model applicability and associated risks against your own business scenarios. For third-party model API services offered via the HUAWEI CLOUD Marketplace, such services

are provided by third-party vendors. The respective rights and obligations for these services are defined between you and the third-party vendor under the relevant agreements. You shall perform necessary assessment on security, accuracy, compliance and business applicability for third-party models and their outputs prior to business adoption. Do not directly apply them to critical business decisions or high-risk operations without validation.

- **Question 6: How is authentication performed for MaaS invocations? How can API Keys be managed securely?**

  Answer: API Key authentication is adopted, with keys generated and managed under HUAWEI CLOUD's unified identity system. For partner resale scenarios, API Keys can be allocated per client with resource isolation and billing reconciliation enabled. Safeguard your credentials, avoid exposure via public channels, and monitoring and key rotation are recommended. In case of anomalies (e.g., 401 errors), contact technical support for recovery.

- **Question 7: How does HUAWEI CLOUD secure the content data and personal data I use within MaaS?**

  Answer: Within HUAWEI CLOUD's control scope, HUAWEI CLOUD delivers corresponding security assurances for MaaS data transmission links, hosted storage and underlying storage environments. Mechanisms including secure transmission, network and tenant isolation, storage encryption, key protection, runtime-environment isolation and auditing mitigate risks of unauthorized data access, tampering or leakage. You still bear corresponding responsibilities for your own data security, including ensuring legitimate data sources, obtaining necessary authorizations, clarifying usage purposes, and implementing data minimization, sensitive-information identification and desensitization, access permission control, encryption and lifecycle management according to data sensitivity. For personal data or other sensitive information contained in your datasets, minimize unnecessary ingestion into models. Where processing is mandatory, implement protective measures such as desensitization before data is fed into MaaS. HUAWEI CLOUD also provides capabilities such as IAM, Data Security Center (DSC), Data Encryption Workshop (DEW), and Cloud Trace Service (CTS) to help you implement data-protection measures within your scope of responsibility.

- **Question 8: Will HUAWEI CLOUD use my personal data, prompts, uploaded data or model outputs for model training or optimization?**

  Answer: In accordance with HUAWEI CLOUD's *MaaS (Model-as-a-Service) Service Statement*, "Your Content" refers to uploaded data, prompts, model outputs and other materials processed by you on the platform. Except for delivering platform-related services to you, HUAWEI CLOUD will not use Your Content (including for model training and optimization) without your explicit authorization. If Your Content contains personal data, it will be handled under the same rules. Scenarios required by laws, regulations and regulatory requirements shall be governed by the service statement and relevant agreements. For specific data-processing rules, mutual rights and obligations and the latest terms, please refer to HUAWEI CLOUD's ***MaaS Service Agreement***.

- **Question 9: Can my data be accessed by other tenants when using a shared resource pool?**

  Answer: HUAWEI CLOUD MaaS isolates data, model assets and computing resources of different tenants via tenant isolation, resource isolation and runtime-environment security mechanisms to mitigate risks of unauthorized cross-tenant access. Sharing underlying resources does not mean that different clients can share or access each other's data. Meanwhile, you shall properly manage your own accounts, API Keys, data-access permissions and security configurations to prevent unauthorized data access caused by misconfigured permissions or credential leakage on the client side. The scope of HUAWEI

CLOUD's processing of client content shall be subject to bilateral agreements, the MaaS service statement, client authorizations and applicable laws and regulations.

- **Question 10: What should I do when an in-use MaaS model undergoes version upgrade, capability adjustment or decommissioning?**

  Answer: Models may be upgraded, have their capabilities adjusted, or be decommissioned due to version iteration, third-party end-of-maintenance, security risks, regulatory requirements or other causes. Subject to the current MaaS Service Statement, for users with invocation records within three months prior to the model decommission date, HUAWEI CLOUD will give advance notice of model decommissioning as stipulated in the statement, in the absence of special legal or regulatory requirements. You shall keep track of announcements released via the HUAWEI CLOUD official website, console and related notifications, and complete model upgrade or business migration within the specified time frame. After switching to a replacement model, re-verify the continued validity of model capabilities, input-output security policies and business controls. When the original service is no longer used, deactivate redundant access credentials and relevant configurations accordingly.

- **Question 11: How should compliance admission, cross-border data transfer and regional**

  Answer: When using MaaS for overseas business, clients shall first confirm available MaaS regions, supported models, service terms and actual data-processing locations in target markets, and assess compliance requirements against business scenarios, data types and locally-applicable laws.

  When performing model inference on this platform, HUAWEI CLOUD will allocate computing resources located within the People's Republic of China and process Your Content within the People's Republic of China, including but not limited to your input prompts and model inference results. Your Content will not be persisted during inference and will be deleted immediately upon inference completion. You are responsible for cross-border compliance of input prompts and associated data, and agree to appoint HUAWEI CLOUD as a data processor to process such data, including input prompts, model inference results and cross-border data-transfer matters. For detailed requirements, please refer to HUAWEI CLOUD's ***MaaS Service Agreement***.

8.3 HUAWEI CLOUD Security and Privacy Services

Clients are responsible for the security of their own data. To safeguard the security of client content, clients may adopt additional security measures aligned with the appropriate security level for their content. Such additional security measures can be provided by either HUAWEI CLOUD or third-party vendors.

HUAWEI CLOUD understands client requirements for security and privacy protection. Drawing on its extensive security-and-privacy-protection practices and technical capabilities, HUAWEI CLOUD offers a portfolio of cloud-based security and privacy-protection services for clients to select from. These cloud services cover products across networking, database, security, management and deployment tools. Built-in capabilities including data protection, data deletion, network isolation, permission management, disaster-recovery backup and security auditing help clients strengthen their security assurances.

- **Security Compliance**

Product Name	Product Introduction	Core Function
Web Application Firewall (WAF)	WAF can detect and defend against website service traffic from multiple dimensions, and work with deep machine learning to intelligently identify malicious request features and defend against unknown threats, preventing common attacks such as SQL injection and cross-site scripting. Clients can use WAF to protect their websites or servers from external attacks, preventing these attacks from affecting the availability and security of web applications or consuming excessive resources, and reducing the risk of data tampering and theft.	Security protection
Large Model Firewall	Fully safeguard the security of AI applications, defend against prompt-based attacks, and ensure content compliance and data privacy. For AI workloads such as intelligent assistants and Agents, it implements security governance over input and output traffic. It supports threat-detection capabilities covering prompt injection, sensitive-information leakage, model jailbreaking, computational-resource DoS, Agent hijacking, and RAG poisoning, so as to guarantee stable operation and data compliance for large-model and Agent-based services.	AI Security Protection Content Security Protection Large-Model Security Protection
Cloud Firewall (CFW)	CFW is a new-generation cloud-native firewall. It provides protection at the Internet and VPC borders on the cloud, including real-time intrusion detection and prevention, global unified access control, full-traffic analysis visualization, log audit, and source tracing analysis. In addition, it supports on-demand elastic capacity expansion, AI-based intelligent defense capabilities, and flexible expansion to meet cloud service changes and expansion requirements, enabling users to quickly and flexibly cope with threats. CFW is a basic service that provides network security protection for user services on the cloud.	Asset protection Access control Online defense
Host Security Service (HSS)	HSS provides functions such as asset management, vulnerability management, baseline check, and intrusion detection, helping enterprises manage host security risks more conveniently, detect and prevent hacker intrusions in real time, and meet graded protection compliance requirements. Clients can use HSS to manage host and container security risks and detect intrusion behaviors such as ransomware, mining, penetration, and escape in real time, meeting graded protection compliance requirements.	Asset management Vulnerability management Baseline check Intrusion detection
Database Security Service (DBSS)	DBSS is an intelligent database security service. Based on the machine learning mechanism and big data analysis technology, DBSS provides functions such as database audit, SQL injection attack detection, and risky operation identification. Clients can use DBSS to detect potential risks and ensure the security of cloud databases.	Security audit
Data Encryption Workshop (DEW)	DEW is a comprehensive cloud data encryption service. It provides functions such as dedicated encryption, key management, and key pair management. Its keys are protected by the hardware security module and integrated	Data encryption

Product Name	Product Introduction	Core Function
	with other HUAWEI CLOUD services. Clients can also use the service to develop their own encryption applications. Clients can use DEW to centrally manage keys throughout the lifecycle to ensure data integrity during storage.	
Anti-DDoS Service (AAD)	AAD is a value-added service that protects Internet servers from being unavailable due to heavy DDoS attacks. Clients can configure AAD IP addresses to divert attack traffic to AAD for cleaning, ensuring stable and reliable origin server services.	Security protection
Data Security Center (DSC)	DSC is a new-generation cloud-native data security platform. It provides basic data security capabilities, such as data classification and grading, data security risk identification, data watermark source tracing, and data anonymization. Clients can use the DSC to integrate the status of each phase of the data security lifecycle and build a cloud service panorama to protect the security of data collection, storage/transmission, use, exchange/destruction.	Data tiering and classification Data anonymization Data watermark
Cloud Bastion Host (CBH)	CBH is a unified 4A security management and control platform developed by HUAWEI CLOUD. It provides enterprises with O&M management services that integrate single sign-on (SSO), unified asset management, multi-terminal access protocols, file transfer, and session collaboration. Clients can use the CBH to remotely operate and maintain cloud hosts, improving the access control security capability, protecting resource O&M and system management security, and reducing the risk of unauthorized intrusion to the system and O&M resources.	Authority management
Cloud Certificate & Manager (CCM)	CMM is a service for issuing and managing certificates throughout the lifecycle. It provides SSL certificate management and private certificate management services. Clients can use CCM to enhance the confidentiality and security of SSL certificates and private certificates, enhance the security of access and transmission channels, and reduce the risk of unauthorized access, access, or theft during data transmission and access.	Certificate management
SecMaster	Based on cloud native security, SecMaster provides comprehensive security information and event management capabilities, such as log collection, security governance, intelligent analysis, situational awareness, and orchestration response, to implement automatic security operations and ensure cloud security.	Situational awareness Secure operation

- **Storage**

Product Name	Product Introduction	Core Function
Cloud Backup and Recovery (CBR)	CBR provides easy-to-use backup services for ECSs, bare metal server, EVS disks, on-premises VMware hypervisors, and local file directories on the cloud. In the event of virus intrusion, incorrect deletion, or software or hardware faults, CBR can	Data backup

Product Name	Product Introduction	Core Function
	restore data to any backup point.	
Volume Backup Service (VBS)	VBS creates permanent online incremental backups for EVS disks, automatically encrypts backup data sent from EVS disks, and restores data to any backup point, improving data availability. VBS reduces the possibility of virus intrusion, incorrect deletion, and software and hardware faults, and ensures data security and reliability. It reduces the risk of unauthorized data modification.	Data backup
Cloud Server Backup Service (CSBS)	CSBS can create consistent online backups for multiple EVS disks on an ECS at the same time. CSBS reduces the possibility of virus intrusion, incorrect deletion, and software and hardware faults, ensures data security, and reduces the risk of unauthorized data modification.	Data backup

- **Management and Supervision**

Product Name	Product Introduction	Core Function
Identity and Access Management (IAM)	Provides identity authentication and permission management functions, manages user accounts, such as employee accounts, system accounts, and applications accounts, and controls the operation permissions of these users on their resources. Users can use IAM to take appropriate measures, such as user management, identity authentication, and fine-grained cloud resource access control, to prevent unauthorized modification of content data.	Authority management
Cloud Trace Service (CTS)	Provides operation records for cloud accounts to implement security analysis, compliance audit, and fault locating. You can configure the CTS object storage service to synchronize operation records to CTS in real time. In this way, operation records can be stored for a longer period of time, ensuring data subjects' right to know and enabling quick search.	Security audit
Cloud Eye Service (CES)	Provides a multi-dimensional monitoring platform for ECSs and bandwidth resources. Users can learn about resource usage and service running status on HUAWEI CLOUD through CES, and respond to alarms in a timely manner to ensure smooth service running.	Security audit
Log Tank Service (LTS)	It provides functions such as log collection, real-time query and storage. Users can leverage logs for real-time decision-making and analysis without additional development, improving log-processing efficiency and facilitating daily operation and maintenance scenarios including real-time log ingestion, query and analysis. Clients can retain operation records of personal information via LTS to safeguard the data subjects' right to know.	Security audit

Product Name	Product Introduction	Core Function
Config	The Config (configuration audit) service provides global resource configuration retrieval, configuration history tracing, and continuous audit and evaluation capabilities based on resource configuration. Clients can use Config to view resource details, relationships between resources, and resource history. Clients can also configure compliance rules to check resource compliance to ensure that cloud resource configuration changes meet expectations.	Security audit

- **Network**

Product Name	Product Introduction	Core Function
Virtual Private Network (VPN)	A VPN is used to establish a convenient, flexible, and ready-to-use IPsec encrypted connection channel between a client's local data center and a HUAWEI CLOUD VPC. Clients can use VPNs to implement flexible, integrated, and scalable hybrid cloud computing environments. In addition, the encryption feature of VPNs improves clients' security protection capabilities.	Secure transmission
Virtual Private Cloud (VPC)	A VPC is an isolated and private virtual network environment on HUAWEI CLOUD. Clients can configure IP address segments, subnets, and security groups in a VPC, and apply for elastic bandwidth and elastic IP addresses to set up service systems. A VPC is a private network on the cloud of clients. All clients are isolated from each other, enhancing data security on the cloud.	Network isolation

